THESIS ON ECONOMICS AND BUSINESS ADMINISTRATION H52

Essays on Household Consumption and Income Underreporting

MERIKE KUKK

TUT
PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
Tallinn School of Economics and Business Administration
Department of Finance and Economics

Dissertation was accepted for the defence of the degree of Doctor of Philosophy in Economics on 14 December 2015.

Supervisor: Professor Karsten Staehr, Ph.D.
Tallinn University of Technology, Estonia

Opponents: Professor Matti Virén, Ph.D.
University of Turku, Finland

Jiri Slacalek, Ph.D.
European Central Bank, Frankfurt, Germany

Defence of the thesis: 26 February 2016

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for any academic degree.

/Merike Kukk/

ISSN 1406-4782
ISBN 978-9949-23-893-4 (publication)
ISBN 978-9949-23-894-1 (PDF)

MAJANDUS H52

Esseed majapidamiste tarbimisest ja sissetulekute alaraporteerimisest

MERIKE KUKK

TTÜ
KIRJASTUS

CONTENTS

LIST OF PUBLICATIONS

I Kukk, Merike, Dmitry Kulikov, and Karsten Staehr (forthcoming). Estimating consumption responses to income shocks of different persistence using self-reported income measures. *Review of Income and Wealth*, DOI: 10.1111/roiw.12163. (ETIS 1.1)

II Kukk, Merike (forthcoming). How did household indebtedness hamper consumption during the recession? Evidence from micro data. *Journal of Comparative Economics,* DOI: http://dx.doi.org/10.1016/j.jce.2015.07.004. (ETIS 1.1)

III Kukk, Merike, and Karsten Staehr (2014). Income underreporting by households with business income. Evidence from Estonia. *Post-Communist Economies*, vol. 26, no. 2, pp. 257-276. (ETIS 1.1)

IV Kukk, Merike, and Karsten Staehr (forthcoming). Identification of households prone to income underreporting: employment status or reported business income? *Public Finance Review*, DOI: 10.1177/1091142115616182. (ETIS 1.1)

Author's Contribution to the Publications:

Article I – The author of the thesis had a leading role in systemising the literature and in running the estimations. The author of the thesis co-wrote the article and acted as the corresponding author in the publishing process.

Article II – The author of the thesis is the sole author of the article.

Article III – The author of the thesis had a leading role in systemising the literature, in running the estimations and in writing the article. The author of the thesis acted as the corresponding author in the publishing process.

Article IV – The author of the thesis had a leading role in systemising the literature, in running the estimations and in writing the article. The author of the thesis acted as the corresponding author in the publishing process.

ACKNOWLEDGEMENTS

First and foremost, I am deeply grateful to my supervisor and teacher Prof. Karsten Staehr for giving me feedback and advice throughout the long process of my doctoral studies. He has devoted endless numbers of his precious hours to discussion of my research. Moreover, he has been very influential in how I think about economics, how I teach and how I conduct research. I deeply appreciate the mental and financial support he has given, which has helped me to finalise my doctoral studies and given me the chance to continue my challenging work in academia.

I thank my co-author of Publication I, Dmitry Kulikov from Eesti Pank. I have learned a lot about econometrics while we were working together on my first publication and I have gained a lot from his comments on my other research papers. I am very glad that I have had the chance to work with the brightest econometrician in Estonia.

I am grateful that Prof. Kadri Männasoo, the head of the Department of Finance and Economics (DFE), and Prof. Kaie Kerem, the former head of DFE, have provided me a good environment at the university during my studies. Prof. Kerem has believed in me from the very beginning of my studies and she has let me teach courses in macroeconomics at undergraduate and master's level.

I am grateful to Tairi Rõõm for letting me work as a visiting researcher in the monetary policy and research department in Eesti Pank. The experience has been very valuable as it has facilitated the creation of a research network for me. I am also grateful to Jaanika Meriküll, Lenno Uusküla and other researchers in Eesti Pank who have always been there to discuss any problems researchers face in their work.

I have been pleased to have had many research scientists around me in the Department of Finance and Economics and I particularly admire my office mate Kaire Põder for her capacity for work.

I thank Robin Hazlehurst for his excellent language editing and his language seminars. I have to admit that I will continue to need his help.

Of course the whole process of doctoral studies would have not been possible without the support from my family, from Nils, Kirke and Kreet. They have accepted that I have had to sit with my textbooks or at a computer late in the evenings, at weekends and during holidays. They have survived when I have been away on the numerous conferences and courses abroad.

INTRODUCTION

On 12 October 2015, the Nobel Prize in economic sciences was awarded to Angus Deaton for his work deepening the understanding of household consumption. His important contribution to economics has been to link individual consumption decisions and outcomes for the whole economy. For economic processes to be explained, it is indeed vital that the economic choices made by households be understood where these determine the division of resources between consumption and saving.

The *purpose* of the thesis is twofold. First, the thesis seeks to extend what is known about household consumption behaviour in terms of household income shocks and the household balance sheet. Second, the understanding of household consumption is used to investigate household tax behaviour. The thesis improves the understanding of household consumption and tax behaviour in the framework of the conventional consumption model.

As can be seen from Figure 1, household consumption represented 68 per cent of GDP in the USA in 2014 and on average 62 per cent in OECD countries. The share has slightly increased over time, suggesting that household consumption is an important building block in aggregate demand.

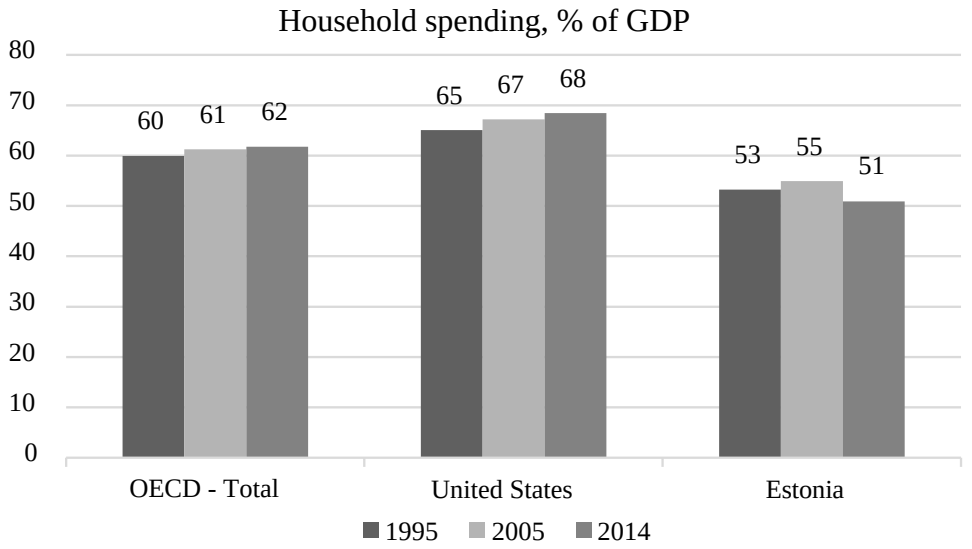


Figure 1. Household spending as a share of GDP.
Source: OECD data on household accounts (available at https://data.oecd.org/)

The large share of consumption in GDP implies that any changes in household consumption entail substantial consequences for the real economy. Figure 2 illustrates the wide fluctuations in household consumption in the last decade, which have contributed markedly to business cycle fluctuations. If policy makers aim to smooth economic fluctuations, it is crucial to understand the causes of the

volatility of consumption. Hall (1986) notes that in earlier equilibrium models households have been considered to respond to exogenous shocks rather than to be a source of business fluctuations. He finds in his work that there are shifts in the consumption schedule, meaning that changes in household consumption behaviour exhibit an important role in business cycles. Therefore it is important to detect factors which induce changes in consumption. Moreover, an understanding of household consumption would help in evaluation of the impact of policy measures on household behaviour.

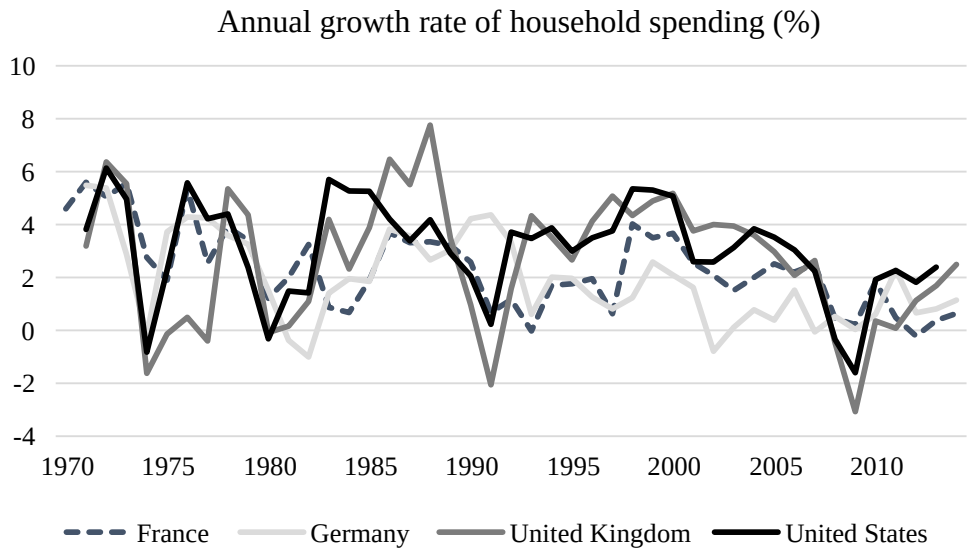


Figure 2. Yearly growth rate of household spending in a set of G7 countries 1970 – 2014. Source: OECD data on household accounts.

Numerous articles investigating household consumption behaviour have been published in recent decades by labour economists, macroeconomists and experts in public economics, the most recent comprehensive overviews being those by Attanasio & Weber (2010), Jappelli & Pistaferri (2010a) and Pistaferri (2015). Although economists know more about household behaviour now than they did some decades ago, there are still a number of unanswered questions. To give an example, Jappelli & Pistaferri (2010a) provide a stock of empirical evidence on the excess sensitivity of consumption to transitory income that is in contradiction to the core of consumption theory. Additionally, Mian & Sufi (2010a) point out that the financial deepening in recent decades has made it necessary to describe in detail the linkages between financial markets and household behaviour.

The first requirement for any understanding of consumption is knowledge of the income components and income processes of households, as human wealth is an important determinant of consumption. The main attention of the literature has

been on the distinction between anticipated and unanticipated income changes or between permanent and transitory income shocks. However, it is important to remember that the information on income that households give to researchers may be incomplete as individuals may conceal a share of their income for various reasons. This makes it vital to recognise which households might be leaving a share of their income unreported and which income components are usually not reported. Furthermore, knowing the relationship between consumption and income allows the extent of income which is not reported to be estimated.

The thesis consists of four published articles; two publications address two research frontiers in the field of household consumption and two publications take this knowledge about household consumption further by investigating household tax behaviour.

The thesis proceeds as follows. Section 1 gives a short review of the consumption theory underlying Publications I and II, while Section 2 presents an introduction to the literature on tax evasion together with the aims and results of Publications III and IV. Section 3 discusses the benefits of the datasets used in the publications. Section 4 reviews plans for further research and Appendices I – IV reprint the four publications.

1. SHORT REVIEW OF THE TOPIC OF CONSUMPTION

1.1. Consumption response to income shocks

The standard workhorse model which is used in macroeconomics to explain consumption and saving behaviour was first introduced by Friedman (1957). His permanent income hypothesis (PIH) postulates that households aim to smooth their consumption over time. The hypothesis follows directly from the solution of the intertemporal maximisation problem faced by each household, where a typical household maximises its expected lifetime utility given the expected discounted stream of all its current and future resources.

The PIH implies that the consumption of an individual household does not depend directly on its current income, but depends on the informational update that current income brings about the future income stream. After receiving new information about a change in its expected discounted current and future income, the household fully adapts its consumption profile to the new circumstances. When new information entails a one-period shift in the income stream with a limited effect on lifetime earnings in contrast, the household absorbs this income change through saving or dissaving, so there is only a limited effect on current consumption. The two types of change to income are commonly known as permanent and transitory income shocks.

A comprehensive review of the methods used to identify income processes is given by Meghir & Pistaferri (2011). An illustration of the implications of income shocks for consumption, like that presented in the working paper version of Publication I in Kukk et al. (2012), can be given here. A permanent income model with certainty equivalence in discrete time with an infinite time horizon is discussed, like in Flavin (1981), Campbell (1987) and Jappelli & Pistaferri (2010a).

The consumption of household i in period t is denoted C_{it} and the exogenous income is Y_{it}. The household chooses its future consumption path in order to maximise the discounted expected utility given a discount rate which is equal to the constant real interest rate r. With a quadratic sub-utility function, the optimal consumption in period t is the share $r/(1+r)$ of the discounted value of current and future incomes as expected by the household in period t:

$$C_{it} = \frac{r}{1+r} \sum_{j=0}^{\infty} \left(\frac{1}{1+r} \right)^{j} \mathrm{E}_{it} Y_{it+j} \tag{1}$$

The term $E_{it}Y_{it+j}$ denotes household i's expected income in period $t+j$ conditional on information available in period t. The solution in equation (1)

entails that the household smoothes consumption across all periods. Consumption only changes between periods when new information about the discounted future income of the household becomes available. If there is a revision of expectations between two time periods, consumption responds as follows:

$$\Delta C_{it} = \frac{r}{1+r}\sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^{j}\left[\mathrm{E}_{it}Y_{it+j} - \mathrm{E}_{it-1}Y_{it+j}\right] \tag{2}$$

The *change* in the consumption of household *i* from period *t - 1* to period *t*, ΔC_{it}, is proportional to the *change* in the discounted sum of expected income in the current and future periods. The impact of income changes on consumption is thus directly linked to their impact on expected future income. Anticipated income changes will have no impact on consumption as they have already been incorporated into the optimal consumption path, while unanticipated income changes will lead to a revision of the optimal path. The effect of unanticipated income shocks on the discounted sum of expected income in the current and future periods, and hence on the optimal choice of consumption, will depend on the persistence of the income process. Shocks to income processes with high persistence will have a larger impact on consumption than shocks to income processes with low persistence.

It is customary to consider AR(1) income processes, where the autoregressive coefficient captures the degree of persistence. It is typically assumed that total income can be decomposed into two AR(1) processes with different autoregressive coefficients.[1] The division into two components follows the framework of permanent and transitory income components pioneered by Friedman.

$$\begin{aligned} Y_{it} &= Y_{it}^{H} + Y_{it}^{L} \\ Y_{it}^{H} &= \rho^{H} Y_{it-1}^{H} + \xi_{it}^{H} \\ Y_{it}^{L} &= \rho^{L} Y_{it-1}^{L} + \xi_{it}^{L} \end{aligned} \tag{3}$$

The total income of household *i* in period *t*, Y_{it}, is the sum of the high-persistence income component Y_{it}^{H} and the low-persistence income component Y_{it}^{L}. The high-persistence component Y_{it}^{H} has the autoregressive coefficient $\rho^{H} \in [0,1]$ and a white noise shock ξ_{it}^{H}. With the widely used assumption that $\rho^{H} = 1$, the high-persistence component is a martingale, implying that income shocks have permanent effects.

[1] In the literature the two components of the income process are often labelled as permanent and transitory. However, as the income components do not need to be truly permanent and transitory the terms high persistence and low persistence are used to differentiate the two income processes.

The low-persistence component Y_{it}^{L} has the autoregressive coefficient $\rho^{L} \in [0,1]$ and the white noise shock ξ_{it}^{L}. With the widely used assumption that $\rho^{L} = 0$, income shocks only affect income in the period in which the shock occurs. The labelling of the two different income components makes it natural to assume that $\rho^{H} > \rho^{L}$.

Expressing equation (2) with the two income processes defined in equation (3) leads to the following equation:

$$\Delta C_{it} = \frac{r}{1+r}\frac{1+r}{1+r-\rho^{H}}\xi_{it}^{H} + \frac{r}{1+r}\frac{1+r}{1+r-\rho^{L}}\xi_{it}^{L} \tag{4}$$

Equation (4) implies that the response of consumption to the income shocks ξ_{it}^{H} and ξ_{it}^{L} depends on the persistence of the corresponding income processes, ρ^{H} and ρ^{L}, and the real interest rate *r*. As it is assumed that $\rho^{H} > \rho^{L}$, the PIH predicts that consumption changes should be more sensitive to a high-persistence shock ξ_{it}^{H} than to a low-persistence shock ξ_{it}^{L}. To assess the quantitative importance of the size of the persistence coefficient, Table 1 shows some examples of consumption sensitivity to income shocks of different persistence computed from equation (4).

Table 1. Theoretical consumption sensitivities to income shocks of different persistence

Persistence coefficient	1.0	0.95	0.9	0.8	0.6	0.4	0.2	0
Theoretical consumption response to income shock	1.0	0.444	0.286	0.167	0.091	0.063	0.048	0.038

Notes: Calculations of eq. (4) using the assumption that the annual real interest rate is 4%.

It follows from Table 1 that consumption sensitivity drops fast according to the PIH when the persistence coefficient falls below 1, as the sum of discounted future earnings is very different depending on whether the discount factor is 1 or smaller. When the persistence coefficient is still as high as 0.9, the theoretical consumption sensitivity is less than 0.3. If the persistence coefficient is 0, the consumption response to the income shock is $r/(1+r)$, meaning the response is negligible given any reasonable value of the real interest rate. Similarly, a shock with an autoregressive coefficient equal to 0.20 dies out very fast, implying that it has a very modest effect on lifetime earnings and consumption.

This example is fairly simple; Kaplan & Violante (2009) provide parameter estimates for the life-cycle version of the standard incomplete markets model, where households have constant relative risk aversion (CRRA) utility. The results are qualitatively similar, as the response to a permanent income shock is

somewhat less than 1 and decreases fast when the persistence of the shock declines.

There is a long history of testing the PIH with empirical data. Hall (1978) was the first to test the orthogonality condition of the PIH, which is that current consumption is the best predictor of future consumption. He uses time-series data for the US and finds that real disposable income does not have any predictive power for consumption. However, it is not possible to derive any conclusions about the PIH from his result because the model did not test the predictable or unpredictable power of *predictable* changes in income, which is the core of the PIH. Campbell & Mankiw (1991) test the PIH by setting up a model which estimates the fraction of consumption that is determined by current consumption, while the rest behaves in accordance with the PIH, meaning consumption responds only to news about permanent income. They find that consumption departs substantially from the PIH.

After the first empirical studies using aggregate data, micro data have been used extensively. Deaton (1987) showed that consumption should be more volatile than income when aggregate data are used, an outcome that is known as the Deaton paradox. He pointed out that household-level data should be used to solve the paradox and to provide a better understanding about the linkage between income and consumption. Zeldes (1989) and Shea (1995) were the first papers, followed by numerous others, to focus on the consumption response to expected income changes or to unanticipated income shocks. Jappelli and Pistaferri (2010a) give a comprehensive review of the main empirical papers. A more recent paper by Gelman et al. (2014) explores the daily consumption reaction to anticipated earnings, while Agarwal & Qian (2014) investigate the consumption response to an anticipated cash payout by the government.

In earlier studies of this topic, a quasi-experimental approach was used to identify income shocks, and Browning & Crossley (2001) and Hryshko et al. (2010) use job loss as an unexpected income change. In the quasi-experimental approach only one income shock is identified and it is not possible to compare the effects of permanent and transitory shocks.

Econometric filtering of permanent and transitory income shocks was introduced by Hall & Mishkin (1982) and extended to the model with CRRA preferences by Blundell et al. (2008). A few papers use the methodology to explore changes in the properties of household income over time (Primiceri & van Rens 2009) or to investigate the consumption response to permanent and transitory income shocks (Jappelli & Pistaferri 2010b). The econometric filtering means that the econometrician uses assumptions about the co-movement of income and consumption to derive the permanent and transitory components from total income. This approach does have some drawbacks when the response of consumption to permanent and transitory components is being investigated, as at the very end it tests whether the initial assumptions which are used to filter the income components hold.

Assessments by households of their own income components can be used to give additional insights about the consumption response to permanent and transitory income shocks. Publication I takes a novel approach to contribute to this field of research. The publication is entitled "Estimating consumption responses to income shocks of different persistence using self-reported income measures", and is co-authored with Dmitry Kulikov and Karsten Staehr. It is one of the first papers to use subjective information received from households to disentangle permanent and transitory income shocks, and it estimates the sensitivity of consumption to those shocks. The paper uses the panel dimension of the Estonian Household Budget Survey (HBS) from 2002 to 2007. The *purpose* of the paper is to compare consumption sensitivity to income shocks of different persistence. To achieve this, the persistence of the two reported income components provided by households is estimated first. The approach differs from other studies as it allows households themselves to distinguish between different income components without imposing any assumption on the persistence of the income components. Other studies look to derive fully persistent and fully transitory income components by assuming that households assess the income process in the same way as an econometrician does.

The more persistent of the self-reported income components is called regular income and the less persistent component is temporary income, emphasising that these do not need to be fully persistent or fully transitory income components. Income shocks are derived from each income component and the consumption responses to income shocks are estimated.

The *results* confirm that there is a different degree of consumption smoothing depending on the persistence of the income shocks. A shock to regular income induces a stronger increase in consumption than a shock to temporary income does. Further estimations reveal, however, that in contrast to the prediction of the theory, consumption also reacts to a *lagged shock* in temporary income. No other study has explored the sensitivity of consumption to a lagged income shock as the standard consumption model assumes that the individual responds to future income shocks, but past income shocks should not affect the individual's consumption. This is the first study to emphasise the importance of including the lagged income shock in the consumption model. The upshot is that the consumption choice is in part affected by factors that cannot be ascribed to the intertemporal smoothing of consumption. The *contribution* of Publication I is a novel identification of income shocks based on self-reported income measures that helps to improve the understanding about the co-movement of consumption and income.

The earlier drafts of the paper were issued in the Eesti Pank Working Paper series in 2012, the paper has been presented in seminars at TTÜ and Eesti Pank, at the Estonian Economic Association Annual Conference (EMS) in 2012, the 4th Conference on Economic Challenges in Enlarged Europe (ECEE4) in 2012, and the International Economic Association (IEA) 17th World Congress in 2014. The final article was accepted for publication in the *Review of Income and Wealth* in October 2014.

1.2. The linkage between household consumption and the balance sheet

Subsection 1.1 focused on the consumption response to changes in lifetime resources induced by income shocks. Alongside human wealth however, non-human wealth can also induce fluctuations in consumption. As utility maximising household behaviour entails that a household consumes all its wealth, both human and non-human, household consumption must also adjust to unexpected changes in real or financial wealth.

Numerous articles investigate the wealth effect, which is the consumption response to real or financial wealth shocks, and these are covered in the literature reviews of recent studies by Browning et al. (2013), Christelis et al. (2014) and Carroll et al. (2011). Studies which investigate the wealth effect on consumption help to explain the consequences for consumption of volatile real estate and stock markets and from this the spill-over into the real economy. The wealth effect has been used to explain the positive correlation between house prices and consumption. Higher house prices increase housing wealth, which in turn increases consumption. It is argued by Muellbauer (2010) and Buiter (2010) that the housing wealth effect is theoretically not that clear as housing value reflects the estimated flow of housing services, which should not be counted as wealth. Nevertheless, most empirical studies find a positive relationship between increases in housing wealth and consumption. Carroll et al. (2011) and Case et al. (2011) compare the wealth effects of the stock market and the housing market and conclude that housing prices have a stronger impact on household consumption than do changes in stock market prices.

The studies exploring the wealth effect on consumption typically neglect the third component of wealth, namely liabilities, or alternatively they use net wealth, which is calculated by deducting a household's liabilities from its assets. Until the 1970s, liabilities comprised a small fraction of households' balance sheets, and so their role was apparently marginal in determining the impact of wealth on consumption, but household debt has grown in most developed countries since the 1980s due to the deregulation of the banking sector. Muellbauer (2010) gives easier credit access as one explanation for the rise in the consumption-to-income ratio in the UK in the 1980s.

Figure 3 shows the rapid increase in the debt-to-disposable income ratio of household sector in a range of advanced and transition countries in Europe. The debt-to-income ratio reached 140 per cent of disposable income in the UK by the mid-2000s while exceeding 100 per cent of disposable income in several European countries. Beck et al. (2008) point out that household debt now exceeds the volume of corporate debt in many countries.

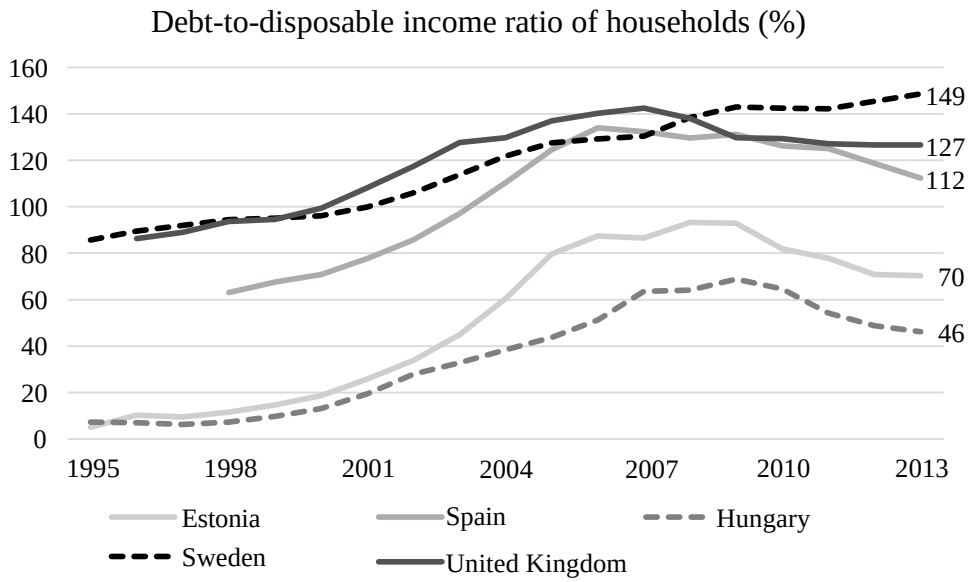


Figure 3. Debt-to-income ratio of households, 1995 – 2014.
Source: Eurostat (available at http://ec.europa.eu/eurostat/data/database, nasa_10_ki)

As a result of the rising role of liabilities in the balance sheet, the composition of net wealth from liabilities, financial assets and real assets varies substantially across households. As noted in Kukk (2014), households with very different levels of liabilities and assets may report the same level of net wealth. Net wealth of 10,000 euros may denote that a household has assets worth 10,000 euros and no liabilities, but it may equally well denote that a household has liabilities of 100,000 euros and assets of 110,000 euros. Therefore it is worth distinguishing between the different components of balance sheets. Moreover, Dynan & Edelberg (2013) argue that there are several channels through which debt affects consumption beyond the standard wealth effect.

The increase in household indebtedness in recent decades has opened a new avenue for academic research. One strand of literature relates the developments in the real estate market to those in the credit markets and investigates the relationship between consumption and housing equity withdrawal. It has been found that in the 2000s households in the US borrowed against their homes to raise their spending (Mian & Sufi 2014a), and Aben et al. (2012) show that the high homeownership rate in Estonia allowed the household sector to extract additional funds for consumption. This channel supported household consumption until the eruption of the 2007-2009 recession.

The recession revealed that household debt may also have adverse effects on consumption. A long list of empirical evidence about the role of household debt in the 2007-2009 recession and in the slow recovery from the recession is

presented in the book "House of Debt" by Mian & Sufi (2014b).[2] Mian & Sufi (2010b) argue that the build-up of excessive household debt before the crises induced a slowdown in economic growth, and high household debt contributed to the severity of the recession. Atif & Mian (2010a) emphasise the need to use micro data to investigate the fundamental causes and consequences of the indebtedness. However, there are only a few papers which investigate the role of household debt in the recession.

Publication II contributes to this area of research. It is entitled "How did household indebtedness hamper consumption during the recession? Evidence from micro data". The *purpose* of Publication II is to assess the linkage between household indebtedness and consumption over the business cycle. Previous studies have focused only on the recession period while Publication II compares the relationship between debt and consumption in different parts of a business cycle. It is known to be the first article in academic literature to cover the whole business cycle when drawing conclusions about the role of debt in a recession.

Publication II undertakes several tasks. First, the article identifies the channels through which indebtedness affects consumption. Second, the article tests whether indebtedness amplified the 2007-2009 recession by investigating whether households held consumption back more during the recession than they did when the economy was growing. Third, it explores how the indebtedness spilled over into consumption in different income groups. Publication II is one of the first in international academic literature to use administrative data in this area of research. The quarterly panel data from 2004:Q4 to 2011:4 are used to estimate a consumption model in which lagged debt-related variables are included.

The *results* of the paper show that there are several channels through which indebtedness affects consumption decisions, as both the lagged debt-to-income ratio and the lagged debt service ratio are negatively related with consumption growth. However, it is the debt service ratio rather than the debt-to-income ratio that captures the amplification of the recession. The results are similar for all income groups, suggesting that the negative relationship between debt and consumption is widespread in the whole population and is not concentrated in any particular sub-group. Publication II provides an important *contribution* to the academic literature as the spill-over of indebtedness to consumption via the debt service burden has not been investigated before.

The findings are valuable for policy suggestions as debt distress measured by the debt service ratio can be alleviated by lower interest rates, which was also observed after the recession erupted. As the majority of mortgages in Estonia have been issued as adjustable interest rate mortgages, the fall in the interest rate did affect most of the indebted households. It can be argued that if the interest rates

[2] The IMF (2014) has announced Prof. A. Mian and Prof. A. Sufi to be among the 25 brightest young economists in the world, indicating that the topic of the role of household debt in the business cycle is on the research frontier for both theoretical and empirical economists.

had not declined, the debt service ratio would have been higher and the consumption drop might have been larger than it actually was.

Earlier draft of the article was issued in TUTECON Working Paper series 5/2014 while a later draft was issued in the Italian Association for the Study of Economic Asymmetries a/working paper 2015/05. The paper has been presented in seminars at TTÜ and Eesti Pank, at the ECEE5 conference in 2013, at the IISES 9th conference in 2014, at the Household Finance and Consumption network meeting in the ECB in 2015 and at the EUA2015 workshop in 2015. The paper received the Vello Vensel award for the best paper of the Doctoral Summer School in 2013. The final article was accepted for publication in the *Journal of Comparative Economics* in July 2015.

2. USING THE CONSUMPTION MODEL TO ESTIMATE INCOME UNDERREPORTING BY HOUSEHOLDS

Knowledge about household consumption can be used beyond the analysis of consumption behaviour as it provides an excellent source for investigating household tax behaviour. Tax evasion and tax avoidance is an ongoing topic in public economics as the efficient collection of taxes is the main base for sustainable public finances. Tax evasion is when an individual does not declare taxable activity, while tax avoidance is when an individual reorganises their economic activity so as to reduce their tax payments (Hindriks & Myles 2012). Tax evasion is part of the shadow economy, which contains all economic activities for which payments are made but not officially declared.

One strand of literature uses theoretical, empirical or experimental methods to investigate the determinants of tax evasion. A comprehensive overview is provided by Alm (2012), who suggests that tax compliance depends on tax rates, audit and fine rates, the tax withholding system, rewards for compliance, and attitudes towards the public sector.

Another strand of the literature aims to estimate the extent of tax evasion, or the tax gap. Several methods for doing this have been used. A direct method uses tax audits where a thorough line-by-line audit is run on a random sample of tax files. Although this method seems to be a straightforward way of estimating the tax gap, Slemrod (2007) points out that line-by-line audits can fail to uncover substantial amounts of non-compliance as the tax authorities do not detect all the concealed income. Another direct method uses surveys and maps non-compliance by questioning individuals or corporations about their tax behaviour. Surveys can be carried out by tax authorities or by third parties. One such survey is by Putnins & Sauka (2011), who use a survey of business managers to estimate the dynamics and the size of the shadow economy in the Baltic countries. The main concern about surveys on tax evasion is that respondents who are hiding their income have every incentive to conceal the truth. The results of direct surveys are very sensitive to the way questions are formulated.

Indirect approaches use various economic and other indicators to extract information about the extent of tax evasion or the unreported economy. First, total economic activity is determined from observation of another economic variable such as an input variable like electricity or a monetary variable like the demand for cash. The gap between total economic activity and the measured activity indicates unreported economic activity. Thomas (1999) points out that indirect models that use macroeconomic indicators are not based on any theoretical models and so a lot of discretion is implied when this method is used. Ahumada, Alvaredo & Canavese (2007) show that studies using the currency approach produce very different results depending on the income elasticity of the demand for currency.

Model-based methods explicitly consider the multiple causes and consequences of the black economy and combine multiple indicators and multiple causes (MIMIC) to estimate the size of unreported economy. The most recent estimates for the world are provided by Buehn & Schneider (2012) using the MIMIC model. Figure 4 provides the estimations for 1999 and 2007, showing that the unreported or shadow economy is twice as large in size in transition countries as it is in high-income OECD countries. The estimations for emerging and developing countries are similar to the estimations for transition countries. However, it is noteworthy that the estimated size of the shadow economy can vary substantially across different specifications of the MIMIC models. In another study by Tafenau, Herwartz & Schneider (2010), somewhat different calculations for transition countries are provided, and the shadow economy in Estonia is estimated to be 17 per cent of reported GDP in 2004, while Buehn & Schneider (2012) estimate it to be 31.1 per cent.[3] Fuest & Riedel (2009) discuss the problems of estimating the total size of tax evasion in developing countries and they note the weaknesses mentioned above. The upshot is that there are some shortcomings to all the approaches which are used to estimate the unreported economy or tax evasion and there is no superior method for producing fully reliable estimates.

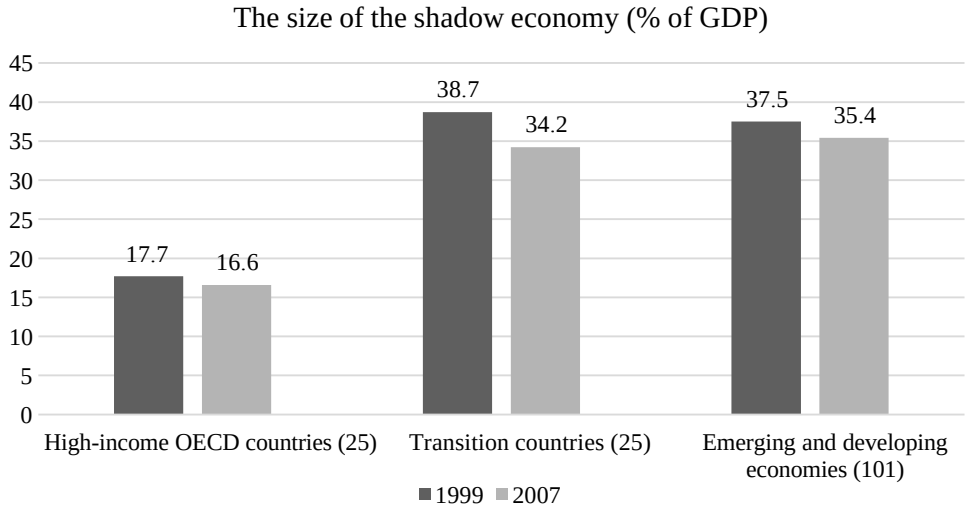


Figure 4. The size of the shadow economy in reported GDP in 1999 and in 2007 using the MIMIC approach.
Source: Buehn & Schneider (2012).

[3] Schneider, Buehn & Montenegro (2010) estimate the size of the shadow economy in Estonia in 2004 to be 40.3 per cent of reported GDP.

Analysis on different sources of tax evasion is needed so that effective policy measures can be formulated that would help to improve tax compliance. The most attention in Estonia has been paid to "envelope wages", which is the term for the payment of wages in cash in order to avoid reporting and taxation (Meriküll & Staehr 2010, Williams 2009). Another important part of the shadow economy has received less attention in the empirical literature. The self-employed are generally better able to evade taxes and some studies even argue that the chance to engage in tax evasion could be one reason for people to choose to be self-employed (Bruce 2000).

Publication III uses the Estonian HBS data indirectly to estimate income underreporting by households. The publication is entitled "Income underreporting by households with business income. Evidence from Estonia", and is co-authored with Karsten Staehr. The *purpose* of the paper is to estimate the income not reported by households with business income in Estonia using the methodology introduced by Pissarides & Weber (1989). The starting point is the Engel curve of food consumption, according to which food consumption and income are closely related for otherwise comparable households. Differences in food consumption behaviour between wage earners and households with business income would then stem from unreported income. The computation of the underreporting factor is based on the estimated coefficients in the Engel curve for food consumption along with variance estimates of the household-specific income for self-employed households and the comparison group.

The methodology has been used in a limited number of papers estimating tax evasion among the self-employed in a few countries, such as the US, the UK, Sweden, Finland, Russia and South-Korea, and a comprehensive literature survey is provided in Section 2 in Publication III. The methodology has not, however, been used on data from any eastern European countries, though these countries have larger shadow economies than western European countries do. These countries have undergone rapid economic restructuring and differ from developed countries in their lower income levels, different economic structures and weaker institutions.

The methodology of Pissarides & Weber (1989) usually gives a range of estimations, as additional assumptions about the *unobservable* permanent income have to be used. Publication III provides exact estimations as the Estonian HBS makes it possible to disentangle transitory and permanent income by using self-reported income information, which is obtained from the results of Publication I.

The baseline *result* is that households with business income leave a more substantial share of their income unreported than do wage earners. The estimated income underreporting does not distinguish between the underreporting of gross income or the over-reporting of business expenses. Publication III has *contributed* to the public discussion about the tax compliance of business income. In recent years the Estonian tax authorities have initiated several moves to limit the practice of shifting personal expenses over to business expenses.

The earlier drafts of the third article were issued in the Eesti Pank Working Paper series in 2013, and the paper has been presented in seminars at TTÜ and Eesti Pank, at the ECEE4 conference in 2012, and at the 2013 European Economic Association conference. The final article has been published in Post-Communist Economies in May 2014. The paper was awarded the title of best paper in social sciences at TTÜ in 2014.

Publication IV continues the topic on income underreporting. It raises a methodological question about the way in which underreporting households are defined. Identifying households which are prone to avoiding taxes is the key for effective policy suggestions. Publication IV is entitled "Identification of households prone to income underreporting: employment status or reported business income?", and is co-authored with Karsten Staehr. It is the first article to explore how the identification of underreporting households affects the estimations of the extent of underreporting.

Studies based on the P&W methodology use two methods for dividing the sample between households which do not underreport their income and households which are assumed to be prone to underreporting income. One method considers the share of reported business income in total reported income in the survey and defines a household as self-employed, and thus prone to underreporting, if the share exceeds a given threshold. The other method uses the household's reported employment status in the survey and defines a household as self-employed if this is the reported employment status.

Pissarides & Weber (1989) is the only study to discuss the importance of the identification choice. They state that they reach similar results when using either of the two methods to identify underreporting, but they only provide the results when self-employment is identified by the share of business income. All the other studies choose one identification method and do not investigate whether the particular choice is of importance for the results. Although it is generally assumed that households which report themselves to be wage earners obtain their main income from their wages, empirical evidence reveals that a substantial number of wage earners earn some business income. Figure 5 illustrates that the share of wage earners receiving some business income varies in European countries from 3 per cent in France to 17 per cent in Greece. One explanation for this is that a second family member might be self-employed, but there is also empirical evidence for second job holding, meaning that wage earners have additional income from business activities and that dual jobholders are self-employed in their second job (Becker 1984, Bregger 1996).

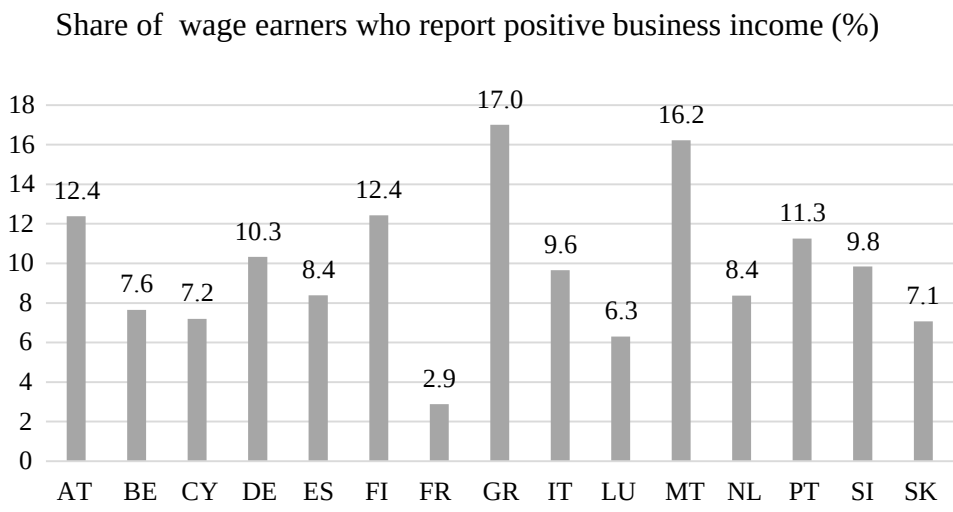


Figure 5. The share of households which report the household head to be a wage earner and which report positive business income.
Notes: The European Union standard country codes are used as country abbreviations.
Source: Household Finance and Consumption Survey, 1st wave from 2009-2010, European National Banks

The *purpose* of Publication IV is to investigate which method identifies households that are prone to underreporting. The article compares the results obtained for Estonia when the two different methods, reported self-employment status and the share of business income, are used, and it finds that the identification method is of material importance. The *results* show that households prone to underreporting can be identified by business income while the reported employment status does not contain any additional information. A substantial share of households prone to income underreporting are overlooked if only households which report themselves to be self-employed are considered to underreport their income. Although this result is derived from the Estonian HBS, it is argued that the results apply to other countries where it is not only the self-employed who earn business income.

The earlier drafts of the fourth article were issued in the TUTECON Working Paper series 1/2014, and the paper has been presented at seminars in TTÜ and Eesti Pank and at the Shadow2013 conference. The final article was accepted for publication in *Public Finance Review* in October 2015.

3. WHY USE DATASETS FROM ESTONIA?

Although all four essays use datasets from Estonia to investigate specific topics, the results of Publications I, II and IV have broader application beyond Estonia. The particular Estonian datasets are used mainly for their unique features which mean they can answer the given research questions and provide a distinctive contribution to the academic literature.

Publication II uses an extensive administrative dataset. It has been shown by Dynan (2012) and Dynan & Edelberg (2013) that survey data have several limitations for investigating the relationship between indebtedness and consumption. The main limitation is the low frequency of survey waves, which are typically run biannually or annually, making it difficult to investigate a specific time period. Additionally, most surveys are cross-sectional, but a panel is needed to observe changes over time. Pistaferri (2015) highlights the problems with existing survey data on consumption and suggests that administrative data are a valuable source for further research on consumption.

The use of frequent administrative data to investigate household consumption has become more widespread recently. Examples include Gelman et al. (2014), who use daily frequency data from a financial aggregation and service application to explore how daily spending responds to anticipated income. Agarwal & Qian (2014) use a proprietary data set obtained from a leading bank in Singapore to investigate the consumption response to an income shock induced by the fiscal policy measure of a cash payout in 2011. Baker (2014) uses household financial data from a large online personal finance website to estimate the impact of debt on the sensitivity of consumption to income shocks. Register data has been used in studies about Scandinavian countries, Andersen et al. (2014) investigate how the loan-to-asset ratio has affected consumption using Danish register data and Yao et al. (2015) explore how a mortgage affects the marginal propensity to consume according to Norwegian register data. Register data do not contain information about consumption and the variable is usually derived as the difference between income and savings.

The dataset which is used in Publication II is very valuable as it contains not only quarterly balance sheet data, but also information on transactions from which income and consumption can be calculated. Comparison of the data with publicly available data confirms that the proxies follow the aggregate income and consumption dynamics. The quarterly panel dataset covers the period from 2004:Q4 to 2011:Q4, traversing the boom, the bust and the recovery. The quarterly data make it possible to estimate the relationship between indebtedness and consumption growth at different times after controlling for income and other variables which may affect consumption change.

Publication I uses the Estonian Household Budget Survey (HBS) from 2002–2007. The dataset contains a self-reported measure of regular income, a feature that is particular to the Estonian HBS. Other surveys in the US or in Europe may contain some information about the regular or usual component of the income, but

in most cases information on income is not collected at such a detailed level as in the Estonian HBS. The Survey of Consumer Finances (SCF) in the US collects data on regular income only when households report their current income to deviate substantially from their usual income. This feature is used by Ackerman & Sabelhaus (2012) in estimating the effect of transitory income shocks on household food spending and explaining the slowdown in the growth rate of food consumption during 2007–2010.

In more recent survey waves of the Estonian HBS the questionnaire has been harmonised with the other European budget surveys and so it now contains very limited data on income.[4] As a result the Estonian HBS from 2002 to 2007 gives the only chance to investigate topics in which it is necessary to identify income components.

The collection of two income measures in the Estonian HBS allows current income to be split into a regular or more persistent income component and a transitory or temporary income component. The split is based on a subjective assessment of income and permits a comparison of how household consumption responds to changes in those subjective income components. The information used by the econometrician is similar to the information used by the household and the symmetric information makes interpretation of the results straightforward.

Publications III and IV use the same features of the Estonian HBS as Publication I does, specifically the detailed information on both food consumption and different sources of disposable income. Household budget surveys in the US and in European countries, such as the US Consumer Expenditure Survey (CEX) and the UK Living Cost and Food Survey (LCF), do contain detailed information on consumption but collect only limited income data.

Moreover, the articles benefit from the availability of the regular income variable when they estimate income underreporting as a point estimate of the extent of underreporting. In most of the other studies a range is calculated for the underreporting factor, as an additional assumption about the covariation between permanent income and underreporting is needed.

Publication IV makes use of another feature in the Estonian HBS, as the survey contains questions that allow different methods to be used to identify households prone to underreporting. Other studies appear to have chosen one identification method for reasons of data availability. Although household budget surveys may contain data on both reported employment status and business income, income data are typically not collected at a detailed level and data on business income may therefore not be available. The UK LCF collects data on weekly spending with diaries but requires households to report only their usual weekly income without detailing the source. The Panel Survey of Income Dynamics (PSID) contains detailed income data but the survey collects food consumption data at a quite general level as no diary is used. Moreover, in some household budget

[4] More precisely, the survey waves after 2008 do not contain detailed data on income, as only one question is asked about regular income.

surveys the time periods of the consumption and the income do not match. An example of this is the US CEX, in which households report annual income for the previous year, while consumption data are collected for the current quarter. In the Estonian HBS the reporting period is the same for consumption and income, so there is no issue about the mismatch between the time periods of the consumption and income data.

The upshot is that it is possible to use datasets from a small country to provide insights with broader application. Although datasets from a small country may seem to limit the scope of research, these datasets may contain unique features which allow a novel approach to be taken to a research question.

4. FINAL COMMENTS

All four essays have contributed to the academic literature in their specific fields. Publication I contributes to the literature which tests the consumption response to permanent and transitory income shocks. The Publication uses a novel identification of income shocks which improves understanding about the co-movement of consumption and income. Publication II identifies different channels through which household indebtedness is related to consumption and finds that the debt distress measured by the debt service ratio is an important channel through which consumption was hampered during the 2007–2009 recession. Publication III finds substantial income underreporting by households with business income in Estonia. The finding contributed to a public debate on tax evasion by the self-employed and the Estonian tax authorities have recently initiated several moves to limit how far personal expenses can be shifted to business expenses. Publication IV finds that the way households prone to underreporting are identified affects the estimated extent of underreporting. The finding highlights that the segment which is prone to underreporting is broader than the segment of the reported self-employed.

Additionally, the papers have opened doors for further research on each of the topics. An important direction for future research is the estimation of different consumption responses to positive and negative shocks to income. Arellano (2014) emphasises the importance of distinguishing between positive and negative income changes as he finds that the persistence of positive and negative income shocks is different, and this also affects consumption in different ways. The importance of investigating the asymmetric response of consumption to income shocks was noted in the working paper version of the first publication (Kukk et al. 2012) and is worth further research.

Publication II will be followed by further research into the linkage between debt repayment problems and consumption. Not only do debt repayment problems affect financial stability, but the effect spills over into the real economy as households adjust their consumption behaviour to cope with financial difficulties. This channel through which household debt spills over into consumption has not previously been investigated.

Finally, no comparative study has been made comparing the extent of income underreporting in a set of countries using the P&W methodology. A comparison of income underreporting across countries would provide further insights into tax behaviour. This research would require a harmonised dataset containing both consumption and income data across a set of countries.

To conclude, the essays in the thesis do not close a research topic but pave the way for further research, and so the Publications of the thesis are just the start of a long journey.

REFERENCES

Aben, M., Kukk, M., & Staehr, K. (2012). Housing equity withdrawal and consumption dynamics in Estonia 2002-2011. *Research in Economics and Business: Central and Eastern Europe*, 4(1), 19-40.

Ackerman, S., & Sabelhaus, J. (2012). The effect of self-reported transitory income shocks on household spending. FEDS Working Paper No. 2012-64.

Agarwal, S., & Qian, W. (2014). Consumption and debt response to unanticipated income shocks: Evidence from a natural experiment in Singapore. Available at http://dx.doi.org/10.2139/ssrn.2245351

Ahumada, H., Alvaredo, F., & Canavese, A. (2007). The monetary method and the size of the shadow economy: A critical assessment. *Review of Income and Wealth*, 53(2), 363-371.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. *International Tax and Public Finance*, 19(1), 54-77.

Andersen, A. L., Duus, C., & Jensen, T.L. (2014). Household debt and consumption during the crisis: evidence from Danish micro data. Danmarks Nationalbank, Working Paper No. 89.

Arellano, M. (2014). Uncertainty, persistence, and heterogeneity: a panel data perspective. *Journal of the European Economic Association*, 12(5), 1127-1153.

Attanasio, O.P., & Weber, G. (2010). Consumption and saving: models of intertemporal allocation and their implications for public policy. *Journal of Economic Literature*, American Economic Association, 48(3), 693-751.

Baker, S. (2014). Debt and the consumption response to household income shocks. Stanford University, Unpublished manuscript available at http://web.stanford.edu/~srbaker/Papers/Baker_DebtConsumption.pdf

Beck, T., Büyükkarabacak, B., Rioja, F., & Valev, N. (2008). Who gets the credit? And does it matter? Household vs. firm lending across countries. World Bank Policy Research Working Paper No. 4661.

Becker, E. H. (1984). Self-employed workers: An update to 1983. *Monthly Labor Review*, 107(7), 14-18.

Blundell, R., Pistaferri, L., & Preston, I. P. (2008). Consumption inequality and partial insurance. *American Economic Review*, 98(5), 1887-1921.

Bregger, J. E. (1996). Measuring self-employment in the United States. *Monthly Labor Review*, 119(3), 3-9.

Browning, M., & Crossley, T. F. (2001). Unemployment insurance benefit levels and consumption changes. *Journal of Public Economics*, 80(1), 1-23.

Browning, M., Gørtz, M., & Leth-Petersen, S. (2013). Housing wealth and consumption: a micro panel study. *The Economic Journal*, 123(568), 401-428.

Bruce, D. (2000). Effects of the United States tax system on transitions into self-employment, *Labour Economics*, 7(5), 545-574.

Buehn, A., & Schneider, F. (2012). Shadow economies around the world: novel insights, accepted knowledge, and new estimates. *International Tax and Public Finance*, 19(1), 139-171.

Buiter, W. H. (2010). Housing wealth isn't wealth. *Economics E-Journal*. 4(22), available at http://www.economics-ejournal.org/economics/journalarticles/2010-22.

Campbell, J. Y. (1987). Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis. *Econometrica*, 55(6), 1249-73.

Campbell, J. Y., & Mankiw, N. G. (1991). The response of consumption to income: a cross-country investigation. European economic review, 35(4), 723-756.

Carroll, C. D., Otsuka, M., & Slacalek, J. (2011). How large are housing and financial wealth effects? A new approach. *Journal of Money, Credit and Banking*, 43(1), 55-79.

Case, K. E., Quigley, J. M., & Shiller, R. J. (2011). Wealth effects revisited 1978-2009. National Bureau of Economic Research, Working paper No. w16848.

Christelis, D., Georgarakos, D., & Jappelli, T. (2014). Wealth shocks, unemployment shocks and consumption in the wake of the great recession. CEPR Discussion Paper No. DP10196.

Deaton, A. (1987) Life-cycle models of consumption: is the evidence consistent with the theory, in *Advances in Econometrics Fifth World Congress*, Vol. 2, ed. T.F. Bewley, Cambridge University Press, Cambridge, 121–48.

Dynan, K. (2012). Is a household debt overhang holding back consumption? *Brookings Papers on Economic Activity*, Spring 2012, 299-362.

Dynan, K., & Edelberg, W. (2013). The relationship between leverage and household spending behavior: evidence from the 2007-2009 survey of consumer finances. Federal Reserve Bank of St. Louis Review, 95 (September/October 2013).

Flavin, M. A. (1981). The adjustment of consumption to changing expectations about future income. *The Journal of Political Economy*, 89(5), 974-1009.

Friedman, M. (1957). *A theory of the consumption function.* Princeton and Oxford: Princeton University Press.

Fuest, C., & Riedel, N. (2009). Tax evasion, tax avoidance and tax expenditures in developing countries: A review of the literature. Report prepared for the UK Department for International Development (DFID).

Gelman, M., Kariv, S., Shapiro, M. D., Silverman, D., & Tadelis, S. (2014). Harnessing naturally occurring data to measure the response of spending to income. *Science*, 345(6193), 212-215.

Hall, R. E. (1986). The role of consumption in economic fluctuations. In The American business cycle: Continuity and change. University of Chicago Press, 237-266

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of Political Economy*, 86(1978), 971-87.

Hall, R. E., & Mishkin, F. (1982). The sensitivity of consumption to transitory income: evidence from PSID households. *Econometrica*, 50(2), 461-481.

Hindriks, J., & Myles, G. D. (2012). Intermediate public economics. 2nd ed. Cambridge: MIT press.

Hryshko, D., Luengo-Prado, M. J., & Sørensen, B. E. (2010). House prices and risk sharing. *Journal of Monetary Economics*, 57(8), 975-987.

IMF (2014). Generation next. 25 economists under 45 who are shaping the way we think about the global economy. In Finance & Development, September 2014, available at https://www.imf.org/external/pubs/ft/fandd/2014/09/pdf/25under45.pdf

Jappelli, T., & Pistaferri, L. (2010a). The consumption response to income changes. *Annual Review of Economics*, 2(1), 479-506.

Jappelli, T., & Pistaferri, L. (2010b). Financial integration and consumption smoothing. *Economic Journal*, 121(553), 678-706.

Kaplan, G., & Violante, G. L. (2009). How much consumption insurance beyond self-insurance? National Bureau of Economic Research, Working Paper Series, No. 15553.

Kukk, M. (2014). Distinguishing the components of household financial wealth: the impact of liabilities on assets in Euro Area countries. Working Papers of Eesti Pank, No. 2.

Kukk, M., Kulikov, D., & Staehr, K. (2012). Consumption sensitivities in Estonia: income shocks of different persistence, Working Papers of Eesti Pank, No. 3/2012

Meghir, C., & Pistaferri, L. (2011). Earnings, consumption and life cycle choices. *Handbook of Labor Economics*, 4, 773-854.

Meriküll, J., & Staehr, K. (2010). Unreported employment and envelope wages in mid-transition: comparing developments and causes in the Baltic countries. *Comparative Economic Studies*, 52(4), 637-670.

Mian, A., & Sufi, A. (2010a). The great recession: Lessons from microeconomic data. *The American Economic Review*, 100(2), 51-56.

Mian, A., & Sufi, A. (2010b). Household leverage and the recession of 2007–09. *IMF Economic Review*, 58(1), 74-117.

Mian, A., & Sufi, A. (2014a). House price gains and US household spending from 2002 to 2006. National Bureau of Economic Research, Working Paper No. w20152.

Mian, A., & Sufi, A. (2014b). *House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening Again.* Chicago: University of Chicago Press.

Muellbauer J. (2010). Household decisions, credit markets and the macroeconomy: Implications for the design of central bank models. BIS Working Paper No. 306

Pissarides, C., & Weber, G. (1989). An expenditure based estimate of Britain's black economy. *Journal of Public Economics*, 39(1), 17-32.

Pistaferri, L. (2015). Household consumption: research questions, measurement issues, and data collection strategies. *Journal of Economic and Social Measurement*, 40(1-4), 97-123.

Primiceri, G. E., & Van Rens, T. (2009). Heterogeneous life-cycle profiles, income risk and consumption inequality. *Journal of Monetary Economics*, 56(1), 20-39.

Putniņš, T. J., & Sauka, A. (2011). Size and determinants of shadow economies in the Baltic States. *Baltic Journal of Economics*, 11(2), 5-25.

Schneider, F., A. Buehn, & Montenegro, C. E. (2010). New estimates for the shadow economies all over the world. *International Economic Journal*, 24(4), 443-461.

Shea, J. (1995). Union contracts and the life-cycle/permanent-income hypothesis. *The American Economic Review*, 85(1), 186-200.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *The journal of economic perspectives*, 21(1), 25-48.

Tafenau, E., Herwartz, H., & Schneider, F. (2010). Regional estimates of the shadow economy in Europe. *International Economic Journal*, 24(4), 629-636.

Thomas, J. (1999). Quantifying the black economy: measurement without theory yet again. *Economic Journal,* 109(456), 381-389.

Williams, C. C. (2009). The prevalence of envelope wages in the Baltic Sea region. *Baltic Journal of Management*, 4(3), 288-300.

Yao, J., Fagereng, A., & Natvik, G. (2015). Housing, debt and the marginal propensity to consume. Available at http://hassler-j.iies.su.se/nordmac/2015/Papers/Jiaxiong.pdf

Zeldes, S. P. (1989). Consumption and liquidity constraints: an empirical investigation. *The Journal of Political Economy*, 97(2), 305-346.

ABSTRACT

The purpose of the thesis "Essays on household consumption and income underreporting" is twofold. First, the thesis seeks to extend what is known about household consumption behaviour in terms of household income shocks and the household balance sheet. Second, the understanding of household consumption is used to investigate household tax behaviour. The thesis improves the understanding of household consumption and tax behaviour in the framework of the conventional consumption model.

The thesis consists of four publications. The first publication is entitled "Estimating consumption responses to income shocks of different persistence using self-reported income measures". The purpose of the paper is to compare consumption sensitivity to income shocks of different persistence. To achieve this, the persistence of the two reported income components provided by households is estimated. Income shocks are derived from each income component and the consumption responses to income shocks are estimated. The results confirm that there is a different degree of consumption smoothing depending on the persistence of the income shocks but consumption also reacts to a lagged shock in temporary income. The contribution of Publication I is a novel identification of income shocks based on self-reported income measures that helps to improve the understanding about the co-movement of consumption and income.

The second publication "How did household indebtedness hamper consumption during the recession? Evidence from micro data" assesses the linkage between household indebtedness and consumption over the business cycle. The results of the paper show that there are several channels through which indebtedness affects consumption decisions. The paper finds that the debt distress measured by the debt service ratio is an important channel through which consumption was hampered during the 2007–2009 recession. Publication II provides an important contribution to the academic literature as the spill-over of indebtedness to consumption via the debt service burden has not been investigated before.

The third publication is entitled "Income underreporting by households with business income. Evidence from Estonia". The purpose of the paper is to estimate the income not reported by households with business income in Estonia using the methodology introduced by Pissarides and Weber (1989). The baseline result is that households with business income leave a more substantial share of their income unreported than do wage earners. Publication III has contributed to the public discussion about the tax compliance of business income. In recent years the Estonian tax authorities have initiated several moves to limit the practice of shifting personal expenses over to business expenses.

The fourth publication "Identification of households prone to income underreporting: employment status or reported business income?" is the first article to explore how the identification of underreporting households affects the estimations of the extent of underreporting. The results show that households

prone to underreporting can be identified by business income while the reported employment status does not contain any additional information. The finding highlights that the segment which is prone to underreporting is broader than the segment of the reported self-employed.

All four publications have contributed to the academic literature in their specific fields. Additionally, the papers have opened doors for further research on each of the topics.

KOKKUVÕTE

Käesolev doktoritöö “Esseed majapidamiste tarbimisest ja sissetulekute alaraporteerimisest” põhineb neljal publitseeritud artiklil. Kahes artiklis keskendutakse majapidamiste tarbimiskäitumise uurimisele ja ülejäänud kahes artiklis kasutatakse teadmisi tarbimiskäitumisest selleks, et uurida majapidamiste maksukäitumist.

Majapidamiste tarbimise arengu uurimine on oluline nii mikromajandusliku heaolu analüüsi kui ka makromajanduspoliitiliste uuringute seisukohalt. Seetõttu on äärmiselt oluline omandada üksikasjalikke teadmisi majapidamiste tarbimiskäitumise sealhulgas majapidamiste tarbimise reaktsiooni kohta sissetulekute muutustele. Doktoritöö esimese publikatsiooni „Tarbimise tundlikkus erineva püsivusega sissetulekušokkidele, kasutades raporteeritud sissetulekukomponente“ eesmärk on uurida, kuidas Eesti majapidamised on reageerinud erineva püsivusega sissetulekušokkidele, kusjuures teoreetilise lähtepunktina kasutati püsiva tulu hüpoteesi (PTH).

Artikli peamine uudsus seisneb Eesti leibkonna eelarve uuringust (LEU) saadud andmete alusel regulaarsete ja ajutiste sissetulekukomponentide eristamises. Selline eristus võimaldab sissetulekukomponentidest eraldada pika kestusega regulaarseid ning lühikese kestusega ajutisi sissetulekušokke. Sellega kaob vajadus määratleda šokke statistilise dekomponeerimise abil, mida on eelnevas teaduskirjanduses kasutatud. Artiklis kasutatud meetod võimaldab ökonometristil kasutada majapidamisega sama informatsiooni.

Artiklis kasutatakse Eesti LEU andmeid ajavahemikul 2002 − 2007, kui LEU sisaldas lisaks detailsele tarbimisinfole ka põhjalikku teavet sissetulekute ja nende tüüpide kohta. Majapidamiste esitatud regulaarsete ja ajutiste sissetulekukomponentide omaduste uurimiseks on kasutatud erinevaid empiirilisi meetodeid, hindamaks ligikaudselt nende püsivust. Analüüsid näitavad, et Eesti leibkonnad ilmselgelt eristavad nimetatud kahe sissetuleku püsivust; šokid regulaarsele tulule on püsivad, samas kui šokid ajutisele tulule on väga vähe püsivad.

Eratarbimise hindamise lähtepunktiks on tarbimise mudel, mis sisaldab regulaarseid sissetulekušokke ning sama perioodi ajutisi sissetulekušokke. Saadud tulemused üldjoontes toetavad PTHd, tarbimise tundlikkus regulaarsete sissetulekušokkide suhtes on oluliselt suurem kui tundlikkus ajutiste šokkide suhtes. Tulemused kehtivad erinevate tarbimisnäitajate puhul, täiendavate kontrollmuutujate lisamisel l ja erinevatel eeldustel sissetulekute püsivuse kohta.

PTH eeldab, et leibkonnad on ratsionaalsed ja tulevikku vaatavad. Selle eelduse järgi ei oma eelmiste perioodide ajutised sissetulekušokid mingit mõju jooksva perioodi tarbimisele. Arvutused aga näitavad, et tarbimine reageerib eelmiste perioodide ajutistele sissetulekušokkidele kui see muutuja on mudelisse kaasatud. Seega ei saa muutust Eesti leibkondade tarbimises buumiaastatel 2002 − 2007 täies mahus selgitada ettevaatava käitumisega, nagu PTH eeldab. Eelmise

perioodi ajutise šoki koefitsient on negatiivne ning väärtuselt ligilähedane jooksva perioodi ajutise sissetulekušoki koefitsiendile.

Üks selgitus, miks tarbimine reageerib muutustele ajutises sissetulekus, on likviidsuspiirangud. Kui majapidamistel ei ole võimalik oma tarbimist likviidsuse puudumise tõttu ühtlustada, siis reageerib tarbimine ka ajutistele sissetulekumuutustele. Täiendavad arvutused näitavad, et see ei ole siiski ainuke selgitus saadud tulemustele, sest ka likviidsuspiiranguteta majapidamiste tarbimine reageerib ajutistele sissetulekušokkidele. Tulemused viitavad majapidamiste proportsionaalsele või niinimetatud peost-suhu tarbimisele. Viimase kohaselt peegeldavad muutused tarbimises lihtsalt muutusi jooksvas sissetulekus.

Uuringus panustatakse tarbimiskäitumise uurimisega seotud akadeemilisse kirjandusse, analüüsitakse tarbimise silumist kiiresti arenevas tärkava turuga majanduses ning kasutatakse majapidamiste poolt esitatud informatsiooni sissetulekušokkide püsivuse kohta. Oluline panus kirjandusse on tulemusel, et majapidamised reageerivad ka eelmiste perioodide sissetulekute muutustele, mida PTH ei eelda.

Doktoritöö teise publikatsiooni „Kuidas pidurdas majanduslanguse ajal majapidamiste laenutase tarbimist? Tulemused mikroandmetest“ eesmärk on selgitada majapidamiste laenude mõju tarbimisele 2008. − 2009. aastate majanduslanguse ajal. Majapidamiste laenude rolli üle reaalmajanduses on viimasel aastakümnel elavalt arutletud, kuna majapidamiste laenumahud on arenenud maades märkimisväärselt suurenenud. Kui varasemal perioodil on majandusteadlased keskendunud ettevõttete laenude ja reaalmajanduse omavaheliste mõjude uurimisele, siis Mian ja Sufi (2010a, 2010b, 2014) tõendavad oma mitmes järjestikuses uurimistöös, et majapidamiste laenud on peamine võti selgitamaks tagasihoidlikku nõudlust viimase majanduslanguse ajal.

Vaatamata elavnenud diskussioonile majapidamiste laenude ümber on sellel teemal ilmunud suhteliselt tagasihoidlik arv teadustöid, sest nende tegemist on peamiselt takistanud sobivate andmebaaside puudumine. Doktoritöö teises artiklis kasutab töö autor kvartaalseid paneelandmeid, mis sisaldavad andmeid indiviidide kohta 2004. aasta lõpust kuni 2011. aasta lõpuni. See on teadaolevalt esimene rahvusvaheline töö, kus kasutatakse majapidamiste tarbimise uurimiseks kvartaalseid andmeid ning võrreldakse majapidamiste laenude ja tarbimise vahelisi seoseid erinevates äritsükli faasides.

Artiklis uurib autor kahte mehhanismi, kuidas majapidamiste laenukoormus võib tarbimist mõjutada. Ühte mehhanismi, nn. laenupohmeluse efekti, kirjeldab laenumahu ja aastase sissetuleku suhe. Teine mehhanism töötab läbi finants-stressi efekti. Seda kirjeldab regulaarsete laenumaksete ja sissetuleku suhe, mis viitab jooksvale laenu teenindamise võimekusele. Neid mehhanisme on siiani kirjanduses käsitletud üpris tagasihoidlikult. Tulemused viitavad, et majapidamiste laenude mõju tarbimisele on kompleksne, kuna laenud mõjutavad tarbimist läbi erinevate kanalite.

Nimetatud kahe laenunäitaja mõju tarbimise muutusele hindab autor tarbimisteooriast lähtuva tarbimismudeliga, kuhu on lisatud mitmed indiviidi majandus- ja finantsolukorda iseloomustavad kontrollnäitajad. Mudelis hinnatakse kvartaalsed koefitsiendid, et identifitseerida laenu mõju erinevate äritsüklite faasides. Juhul, kui eelnevalt võetud laenud võimendavad majanduslangust, siis peaks laenunäitajate negatiivne seos tarbimise muutusega olema majanduslanguse ajal tugevam. Seda hüpoteesi ei ole teadaolevalt üheski teadustöös varem kontrollitud.

Arvutused näitavad, et kahe laenunäitaja mõju tarbimisele on ajas erinev. Laenumahu-sissetuleku suhte mõju tarbimise muutusele on terve perioodi jooksul negatiivne, kuid suhteliselt stabiilne ja näitab mõnevõrra nõrgenevat seost vaadeldava perioodi jooksul. Ilmselgelt ei ole laenu-sissetuleku suhte mõju tarbimisele seotud äritsükliga.

Samas on laenu teenindamise koormuse mõju tugevam 2008. – 2009. aastatel, võrreldes sellele eelneva ja järgneva perioodiga. Teadustöö tulemustest saab järeldada, et laenu teenindamise koormus võimendab majanduslangust läbi tarbimise languse – majapidamised piiravad laenude teenindamise koormuse tõttu oma tarbimist majanduslanguse ajal rohkem kui majandustõusu keskkonnas; ka siis, kui laenu teenindamise koormus ei muutu. Kuid kuna ka laenu teenindamise koormus sel ajavahemikul tõusis, võimendas laenumakse koormuse suurenemine omakorda negatiivset mõju tarbimisele.

Uurides majapidamiste laenude negatiivse mõju jaotust erinevates sissetulekugruppides näitavad arvutused, et tulemused on sissetulekugruppides suhteliselt sarnased. Siit saab järeldada, et laenuteenindamise koormuse võimendav mõju majanduslangusele ei ole koondunud väiksema sissetulekuga majapidamiste segmenti, vaid see hõlmab kõiki sissetulekugruppe.

Kokkuvõtteks saab artikli põhjal väita, et majapidamiste laenudel on äritsüklis oluline roll, kuna laenud võimendavad majanduslangust ning mehhanism toimub eelkõige läbi laenu teenindamise koormuse ehk laenu -stressi kanali. Laenumahu-sissetuleku suhe ei ole majapidamiste poolt nii otseselt tunnetatav kui laenu teenindamise koormus. Viimasele laenunäitajale on rahvusvahelises diskussioonis laenude ja tarbimise üle võrreldes teiste laenunäitajatega vähem tähelepanu pööratud. Artikli tulemused viitavad, et laenu teenindamise koormuse näitaja sisaldab olulist informatsiooni majapidamiste tarbimiskäitumise kohta, sest selgitab seost majapidamiste laenude ja tarbimise vahel.

Artikli tulemused aitavad mõista majapidamiste laenude mõju tarbimisele ka teistes maades, eelkõige Kesk- ja Ida- Euroopa riikides, kus 2000. aastate majandustõusuga kaasnes ka kiire majapidamiste laenumahtude kasv. Artikli tulemused võimaldavad poliitikategijatel teha asjakohaseid soovitusi.. Laenu teenindamise koormuse negatiivset mõju tarbimisele saab leevendada intressipoliitikaga. Eestis moodustavad laenumahust suurima osakaalu ujuva intressiga eluasemelaenud ning seetõttu on intressimäärade muutmisega võimalik mõjutada otseselt laenu teenindamise koormust ja seeläbi tarbimise muutust. Laenude intressimäärad on alates 2008. aastas lõpust langenud ja eeldatavalt on

seetõttu laenukoormuse negatiivne mõju tarbimisele olnud väiksem kui oleks olnud ilma intressmäärade muutumiseta.

Doktoritöö kolmandas publikatsioonis „Ettevõtlustulu teenivate majapidamiste sissetulekute osaline esitamata jätmine Eestis“ kasutab autor tarbimismudelit, et uurida majapidamiste maksukäitumist. Paljud Eesti majapidamised teenivad osa oma kogusissetulekust (üksik)ettevõtlusest või muust individuaalsest töisest tegevusest. Kuna sellise tulu deklareerimine on suuremalt jaolt rahateenija enda ülesanne, on suhteliselt lihtne jätta osa oma sissetulekust esitamata. Samal põhjusel on sissetuleku varjamist suhteliselt keeruline avastada. Seetõttu keskendutakse paljudes maksudest kõrvalehoidumist käsitlevates uuringutes ümbrikupalga saajatele ja ainult vähestes uuringutes tegeletakse ettevõtlustulu teenivate majapidamistega. Doktoritöö kolmandas artiklis esitatud tulemused täidavad selle lünga Eestis.

Artiklis on kasutatud Pissaridese & Weberi (1989) poolt kasutusele võetud meetodit (P&W), mille abil arvutatakse majapidamiste eelarveküsitluse põhjal välja see osa ettevõtlustulu teenivate majapidamiste sissetulekust, mis jäetakse küsitluses esitamata. Meetodi järgi eeldatakse, et toidu tarbimine ja sissetulek on omavahel tihedalt seotud, s.t toidukulutuste osakaal sissetulekus on sarnastes majapidamistes samaväärsel tasemel. See eeldus võimaldab võrrelda sissetulekuid ja toidukulutusi kahe grupi vahel: majapidamised, kus peamise sissetuleku annab palgatöö ja majapidamised, kus tulu saadakse ka ettevõtlusest. Ettevõtlustulu saavate majapidamiste keskmine kogusissetulek on üldjuhul küsitluste järgi väiksem ja nende keskmine toidutarbimine on suurem kui palgatööga teenivate majapidamiste vastavad keskmised näitajad. Kuna kulutusi toidule kajastavad vastajad küsitlustes tavapäraselt täpsemalt kui oma sissetulekuid, võib arvata, et ettevõtlustulu teenivad majapidamised mitte ei kuluta toidule rohkem, vaid näitavad tegelikust väiksemat sissetulekut.

Kuigi arvutused tehakse majapidamiste eelarveküsitluse põhjal, võib eeldada, et inimesed on oma andmete avaldamises erinevatele institutsioonidele järjepidevad ning teatavad sama suure sissetuleku ka teistele asutustele. Kui see eeldus paika peab, siis annavad uuringu arvutused aimu ka sellest, kui suure osa oma tuludest jätavad ettevõtlustulu teenivad majapidamised maksuametile deklareerimata.

Uurimuses kasutatakse LEU andmeid aastatest 2002 − 2007 ning arvutatakse välja, kui suure osa oma sissetulekust jätavad ettevõtlustulu teenivad majapidamised esitamata võrreldes palgatööst elatuvate majapidamistega. Uurimuses on P&W meetodit mõnevõrra muudetud ja jooksva sissetuleku asemel on kasutatud majapidamiste regulaarset sissetulekut. Majapidamiste toidutarbimine on ühtlasem kui sissetulek, sest majapidamise tarbimine on seotud regulaarse või püsiva sissetulekuga, mida sissetuleku ajutine tõus või langus ei mõjuta. P&W meetod kasutab jooksvat, mitte püsivat sissetulekut, mistõttu on arvutused ebatäpsemad ja annavad suurema vahemiku sissetulekute avaldamata jätmise suurusjärgu kohta. Kasutades Eesti LEU-d, kus majapidamised avaldavad

ka oma regulaarse sissetuleku, on sissetuleku varjamise ulatust võimalik arvutada täpsemalt.

Majapidamised jaotatakse võrdlemise eesmärgil kahte gruppi. Esimesse kuuluvad majapidamised, kus ettevõtlustulu annab vähemalt 20% avalikustatud kogutulust ja majapidamised teise, kus sissetulek ei sisalda ettevõtlustulu, s.t nende põhitulu tuleb palgatulust. Ettevõtlustulu hulka arvestatakse tulu registreeritud individuaalsest tegevusest, tulu eraviisilisest tasuliste teenuste osutamisest, tulu kaupade tootmisest, kaupade või teenuste vahendusest saadud tulu või muu äritegevusest saadud tulu.

Arvutused näitavad, et majapidamised, mis teenivad ettevõtlustulu, on jätnud leibkonnaküsitluses esitamata ca 60% oma tegelikult teenitud tulust. Tegeliku tulu all peetakse siin silmas palgatööst elatuvate majapidamistega võrdväärset sissetulekut. Sissetulekute avalikustamisel väiksemas mahus võib olla mitu põhjust: kas osa sissetulekust varjatakse või esitatud tuludest on maha arvatud suurem summa kulutusi, kuna osa isiklikke kulutusi esitatakse ärikulutustena. Isegi kui toidutarbimise mõningane erinevus on tingitud erinevatest eelistustest, on vahe palgatööliste ja ettevõtlustulu teenivate majapidamiste vahel märkimisväärne ning viitab sissetulekute osalisele avalikustamata jätmisele.

Võrreldes maadega, mille kohta on samuti olemas P & W meetodiga tehtud arvutused ettevõtlustulu teenivate majapidamiste sissetulekute esitamata jätmisest, on tulemused Eesti majapidamiste kohta märkimisväärselt suuremad. USA, Kanada, Ühendkuningriigi, Rootsi ja Soome puhul on arvutatud, et ettevõtlusega tegelevad majapidamised jätavad avalikustamata umbes 30 – 40% oma sissetulekutest. Samas on tulemused sarnased teiste uuringutega ümbrikupalkade ja üleüldise varimajanduse kohta, mis viitavad sellele, et üleminekumajandusega riikides on varjatud tegevused levinud kordades suuremas ulatuses kui arenenud maades.

Doktoritöö neljas publikatsioon on „Sissetulekute alaraporteerimisele kalduvate majapidamiste identifitseerimine: kas tööalane staatus või raporteeritud ettevõtlustulu?“. Artikli eesmärk on võrrelda sissetulekute alaraporteerimise ulatust erinevate identifitseerimisviiside korral ja välja selgitada, kuidas määratleda kõige täpsemalt alaraporteerimisele kalduvaid majapidamisi. Uurimistöödes, kus kasutatakse P&W meetodit, on kasutatud kahte erinevat identifitseerimisviisi. Ühes osas uurimistöödest kasutatakse perekonnapea ütlusi selle kohta, kas ta on palgatööline või (üksik)ettevõtja. Teistes defineeritakse uuringu huviorbiidis olevad majapidamised selle järgi, kui suur on nende äritulu majapidamise raporteeritud sissetulekus. Pissarides & Weber (1989) ainult mainivad võimalust kahel viisil identifitseerida ja näitavad oma artiklis ainult tulemusi, defineerides alaraporteerimisele kalduvad majapidamised ettevõtlustulu osakaalu abil. Teiste maade kohta tehtud uurimustes, kus kasutatakse sissetulekute alaraporteerimise arvutamiseks P&W meetodit, ei eristata ega võrrelda ettevõtlustulu teenivate majapidamiste erinevaid määratlusi.

Eestis kahe erineva definitsiooni järgi määratletud majapidamised üks-ühele ei kattu. Eesti Leibkonnauuringu järgi ca 1/3 majapidamistest, kus perepea

määratleb end iseenda tööandjana, ei avalikusta ettevõtlustulu; ligi pooled majapidamistest, kus perepea määratleb end palgatööliseks, raporteerib ka ettevõtlustulu. Arvutused näitavad, et erinevate identifitseerimismeetodite kasutamine viib erinevate tulemusteni. Kui määratleda alaraporteerimisele kalduvad majapidamised tööalase staatuse järgi, siis ettevõtjana raporteeriva perepeaga majapidamine jätab avalikustamata 28 protsenti oma sissetulekust võrreldes palgasaajaga perepeaga majapidamisega. Kui võrrelda ettevõtlustulu saavaid majapidamisi ettevõtlustulu mittesaavate majapidamistega, siis esimesed jätavad raporteerimata 43 protsenti oma kogusissetulekust.[5] Erinevad tulemused on tingitud sellest, et arvutustest selgub alaraporteerimine baasgrupi suhtes ja lisaks võrdlusgrupile on erinevate määratluste korral ka baasgrupp erinev.

Erinevuste täpsemaks uurimiseks eristati mudelis majapidamised kahes identifitseerimise lõikes. Baasgrupiks võeti majapidamised, mille perepea on palgasaaja ja majapidamine ei saa ettevõtlustulu. Ülejäänud majapidamised jagati kolme rühma: 1) majapidamised, mille perepea on palgasaaja, kuid majapidamised teenivad ka ettevõtlustulu; 2) majapidamised, mille perepea on ettevõtja ja raporteerib ettevõtlustulu; 3) majapidamised, mille perepea on ettevõtja, kuid ei raporteeri ettevõtlustulu. Mudelis arvutati sissetulekute alaraporteerimine eraldi kõigi kolme grupi kohta. Selleks, et arvutused oleksid kahe ettevõtlustulu saava grupi vahel võrreldavad, viidi ettevõtlustulu osakaalu jaotus raporteeritud tulust erinevates gruppides üksteisega vastavusse (nn. matching). Arvutustest ilmnes, et nii palgasaajad kui ettevõtjad, kes raporteerivad ettevõtlustulu, jätavad raporteerimata üle poole oma kogutulust (53% – 56%). Samas ettevõtjad, kes ei raporteeri ettevõtlustulu, ei näi oma tulu varjavat.

Arvutused näitavad, et majapidamiste määratlemine ettevõtlustulu järgi annab usaldusväärsemaid andmeid sissetulekute esitamata jätmise kohta, sest ettevõtlustuluga palgasaajad käituvad samuti kui ettevõtlustuluga ettevõtjad. Kui jätta see fakt arvestamata ja käsitleda ettevõtlustuluga palgasaajaid sarnaselt ettevõtlustuluta palgasaajatega, siis jääb oluline osa sissetulekute alaraporteerimisest kaardistamata. See on oluline järeldus ka teiste tööde jaoks, kus kasutatakse P&W meetodit sissetulekute alaraporteerimise ulatuse hindamiseks. Samuti aitab see poliitikategijatel suunata oma meetmed mitte niivõrd ettevõtjatele kui ettevõtlustulule, sõltumata kas see on majapidamiste põhitulu või kõrvaltulu.

Doktoritöö artiklid katavad kolme uurimisvaldkonda majapidamiste tarbimiskäitumises ja maksukäitumises ning saadud tulemused annavad olulise panuse teaduskirjandusse. Artiklites esitatud uurimistulemused võimaldavad antud valdkonnaga seotud temaatikat veelgi sügavamalt uurida. Seetõttu ei käsitle autor oma doktoritööd kui ühe uurimisteema lõpptulemust, vaid kui pika tee algust.

[5] Tulemus on erinev kolmanda publikatsiooni arvutustest, kuna siin ei ole välja jäetud majapidamisi, kus ettevõtlustulu on positiivne, kuid väiksem kui 20 protsenti kogutulust. Võrdluse jaoks neljandas publikatsioonis on oluline kaasata arvutustes kogu valim.

Appendix 1. Publication I

Reprint from

Kukk, Merike, Kulikov, Dmitry, and Karsten Staehr (forthcoming). Estimating consumption responses to income shocks of different persistence using self-reported income measures. *Review of Income and Wealth*, available online from 7 January 2015 at http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1475-4991, DOI: 10.1111/roiw.12163

the review of income and wealth

Review of Income and Wealth 2015
DOI: 10.1111/roiw.12163

ESTIMATING CONSUMPTION RESPONSES TO INCOME SHOCKS OF DIFFERENT PERSISTENCE USING SELF-REPORTED INCOME MEASURES

BY MERIKE KUKK*

Tallinn University of Technology, Estonia

DMITRY KULIKOV

Eesti Pank, Estonia

AND

KARSTEN STAEHR

Tallinn University of Technology and Eesti Pank, Estonia

Models of intertemporal consumption choice posit that consumption reacts more strongly to income shocks with persistent effects than to shocks with temporary effects. This prediction is tested using data from the Estonian Household Budget Surveys for 2002–07. Questions in the survey make it possible to distinguish between two income components of different persistence, using the individual households' subjective income classification. Estimations confirm that households distinguish income components of different persistence and react to these differently; the consumption response to income shocks with persistent effects is significantly higher than the response to shocks with only temporary effects. Further analysis reveals, however, that consumption also reacts to lagged shocks to temporary income even when the households are not liquidity constrained, suggesting that their behavior is not fully consistent with the standard forward-looking unconstrained consumption models.

JEL Codes: D12, D91, E21, R22

Keywords: consumption smoothing, excess smoothness, income persistence, income shock, intertemporal consumption allocation, self-insurance

> [A]lthough the agent may be able to discriminate between a transitory and a permanent shock, the econometrician is not. As a result, econometric identification of separate income shock components is difficult in the extreme. (Pistaferri, 2001, p. 465)

1. INTRODUCTION

This paper studies the response of household consumption to shocks in income processes with different persistence, using data from a panel of Estonian

Note: The authors would like to thank Arvo Valtin from Statistics Estonia for help on numerous data issues, the seminar participants at Eesti Pank and Tallinn University of Technology, participants of EMS 2012, the ECEE4 and the IEA 17th World Congress, as well as two anonymous referees for useful comments. Kukk and Staehr acknowledge support from Estonian Target Financing grant no. SF0140059s12. All errors remain the responsibility of the authors. The views expressed are those of the authors and not necessarily those of Eesti Pank.

*Correspondence to: Merike Kukk, Department of Finance and Economics, Tallinn University of Technology, Akadeemia tee 3-478, 12618 Tallinn, Estonia (merike.kukk@ttu.ee).

households for the period 2002–07. The dataset makes possible a decomposition of household income into a process that households assess as persistent and a process that is assessed as exhibiting little or no persistence. Unlike most previous studies, this identification of income processes of different persistence is based on the reporting of the households and does not rely on the theory-based restrictions that are otherwise required.

Following the seminal contribution of the Permanent Income Hypothesis (PIH) by Friedman (1957), models of intertemporal consumption choice typically predict that households respond differently to shocks of different persistence. Households respond one-to-one to income shocks that are expected to have a permanent impact on future income, while the consumption response to shocks with transitory effect is negligible. Numerous studies have sought to provide estimates of consumption responses to different types of income shock in order to gain a better understanding of the factors that affect consumption.

It is challenging to devise empirical tests of the intertemporal consumption model, particularly tests that can estimate the consumption response to shocks in income processes of different persistence. The main reason for this is that when an income change is observed for a household, the amount of supplementary information available on the persistence of the income change is usually limited. The decomposition of income into a persistent and a transitory component is needed not only for investigation of the consumption choices of households, but also for other purposes like analysis of earnings inequality, as in Gustavsson (2008). Three different ways to decompose income shocks into components of different persistence have been devised in the literature.

One approach is to use quasi-experimental data in which specific episodes of income changes can be classified according to their expected impact on the income path of the households. Another approach relies on model-based or statistical decomposition of observed income shocks into permanent and transitory components. Such econometric identification requires either very long time series or additional restrictive assumptions about the co-movement of consumption and income (cf. also the citation above from Pistaferri, 2001). Finally, it might be possible to deduce the degree of income shock persistence directly from information provided by the individual household. The use of subjective or self-reported income measures to identify the persistence of income shocks is very rare in the literature as the required information is typically not available in household surveys.

This paper uses the latter approach as the Estonian Household Budget Survey (HBS) uniquely permits a decomposition of household income into two distinct components: one for which the household expects changes to have long-lasting effects, and one for which the household expects changes to be transitory. Using supplementary statistical evidence, we find that Estonian households do indeed split their income according to this classification as the dynamics of the two income processes are markedly different. The decomposition enables us to study the consumption response of Estonian households to income shocks of different persistence.

The paper also offers a view on intertemporal consumption choice at the household level in an environment characterized by a high degree of macroeco-

nomic uncertainty. The Estonian households in this study experienced a rapid increase in their income throughout the period 2002–07, while high inflation, booming real estate prices, and increased foreign indebtedness contributed to increased uncertainty (Brixiova *et al.*, 2010). This macroeconomic environment offers an exciting setting for testing the predictions of consumption models built on a forward-looking optimizing behavior of rational agents.

The rest of the paper is organized as follows. Section 2 provides a brief overview of the theoretical and empirical literature on consumption response to income changes. Section 3 introduces the Estonian HBS data and examines the properties of the income shocks that are identified using the ancillary survey information. Section 4 presents the results of the consumption estimation using the income shocks identified in Section 3. Finally, Section 5 summarizes the empirical findings.

2. Consumption Response to Income Changes: A Brief Literature Overview

Theories of intertemporal consumption assume that the household plans its consumption choices over present and future periods simultaneously, while seeking to maximize expected aggregate utility. Such models typically predict that a forward-looking rational household will smooth consumption over time given its expected future income stream (Jappelli and Pistaferri, 2010a). The reaction of consumption to income shocks will therefore depend on the impact of the shocks on the sum of current and expected future income. The consumption of an individual household does not depend directly on its current income, but on the information contents of the current income for the future income stream.

The current income provides information that allows the household to update its expectation of future income, and the household consequently adapts its consumption profile to the new information. One example is the Permanent Income Hypothesis which asserts that if an unanticipated change in income is deemed to be fully persistent, the household will alter consumption one-to-one. On the other hand, a transitory shock depicting an unanticipated one-period shift in the income stream implies a limited effect on lifetime earnings; the household will absorb most of the income change through saving or dissaving and the effect on current consumption will be limited, depending on the remaining lifespan.

It is clear that many factors affect the response of consumption to income shocks or, phrased differently, the degree of insurance of consumption against income shocks (Blundell *et al.*, 2008). Besides the persistence of the shock and whether or not it is anticipated, factors such as the household's time and risk preferences, liquidity constraints, and information constraints will be of importance (see the long list in Blundell *et al.*, 2008, including footnote 3).

The theoretical effects of the persistence of the income shock have been examined in a number of studies. Kaplan and Violante (2010) and Hryshko (2013) compute the theoretical consumption response to an unanticipated income shock

given different assumptions about the model of intertemporal consumption choice and the persistence of the shocks. They show that the consumption response decreases fast when persistence falls below 1. Kaplan and Violante (2010) present that the consumption response is expected to be below 0.5 when the persistence is 0.93. Kukk *et al.* (2012) tabulate the consumption response to income shocks in the baseline PIH model for a wider range of persistence. They show that when the persistence coefficient is 0.8, the theoretical consumption response is below 0.2. However, theoretical models of rational forward-looking choice predict a stronger consumption response to persistent or permanent unanticipated income shocks than to temporary or transitory unanticipated income shocks.

Empirical testing of the intertemporal consumption model is typically based on a classification of unanticipated income shocks into different components according to their impact on the household's expected discounted income. Jappelli and Pistaferri (2010a) provide a comprehensive review of the empirical literature using a range of approaches to disentangle observed income shocks into anticipated and unanticipated changes. To test the hypotheses of the intertemporal consumption model, the unanticipated income shocks are further divided into permanent and transitory shocks (Chao, 2003).

Studies investigating income dynamics typically decompose the income into a highly persistent component and a transitory or less persistent component (Heathcote *et al.*, 2009). We use the term highly persistent (superscript H) and less persistent (superscript L) to differentiate the two income processes:

$$\begin{aligned} \log Y_{it}^{H} &= \rho^{H} \log Y_{it-1}^{H} + \xi_{it}^{H} \\ \log Y_{it}^{L} &= \rho^{L} \log Y_{it-1}^{L} + \xi_{it}^{L} \end{aligned}. \tag{1}$$

Subscript i indexes the household and t the time. The highly persistent income component $\log Y_{it}^{H}$ is described by its autoregressive coefficient $\rho^{H} \in [0, 1]$ and a white noise shock ξ_{it}^{H}. Under the widely used assumption that $\rho^{H} = 1$, the highly persistent component is a martingale, implying that shocks have permanent effects. The less persistent income component $\log Y_{it}^{L}$ is described by its autoregressive coefficient $\rho^{L} \in [0, 1]$ and the white noise shock ξ_{it}^{L}. Under the widely used assumption that $\rho^{L} = 0$, income shocks only affect the income in the period in which the shock occurs. The labeling of the two different income components implies that $\rho^{H} > \rho^{L}$.

The standard empirical model for estimating the consumption response to shocks in the two income processes in equation (1) takes the form:

$$\Delta \log C_{it} = Z_{it}'\gamma + \beta_1 \xi_{it}^{H} + \beta_2 \xi_{it}^{L} + \varepsilon_{it}. \tag{2}$$

The dependent variable $\Delta \log C_{it}$ is the change in log consumption. The variables in the column vector Z_{it} are preference shifters that affect consumption and γ is a vector of corresponding coefficients. The coefficients β_1 and β_2 capture the response of consumption to shocks to the highly and less persistent income components, cf. equation (1). Finally, ε_{it} is an error term. The model is used in numerous empirical studies (cf. Blundell *et al.*, 2008; Jappelli and Pistaferri, 2010a; Hryshko, 2013).

In *quasi-experimental* data the reaction of households to specific episodes of income changes is used to disentangle the effect of shocks with permanent and transitory effects (see, e.g., Hryshko *et al.*, 2010; Sahm *et al.*, 2010; Aguila *et al.*, 2011). In some episodes the observed income changes are regarded as transitory, such as temporary unemployment, one-off tax refunds, or weather shocks, while in other episodes they are viewed as long-lasting, such as major health problems or retirement. A drawback of quasi-experiments is that they tend to focus on one specific income change, whereas households are likely to be subjected to several shocks at the same time and their reactions to a particular episode may be affected by other income innovations of a potentially different kind.

A model-based *econometric decomposition* of the observed income changes into permanent and transitory components was pioneered in the seminal paper of Hall and Mishkin (1982). They derive from the Permanent Income Hypothesis a set of variance–covariance restrictions between changes in consumption and changes in income across different time periods, and furthermore they impose the assumption that the household's income consists of two parts: a difference-stationary component with innovations that persist indefinitely; and a covariance-stationary component whose innovations dissipate over time. Hall and Mishkin (1982) use data from the U.S. Panel Study of Income Dynamics (PSID) and a derived set of restrictions to estimate the response of household consumption to these two types of innovations. They find that 80 percent of the households consume a constant share of their expected discounted current and future income streams, while the remaining 20 percent follow the *rule-of-thumb* proportional consumption model.

Quah (1990) uses the same model-based decomposition of the income process to derive testable implications about the *aggregate* income and consumption processes. In particular, he relies on the spectral density properties of income changes under the assumption of one difference-stationary component and one covariance-stationary component to argue that smoothness of the observed aggregate consumption process is not inconsistent with considerably more volatile aggregate income, as previously claimed in *Deaton's paradox* by Deaton (1987).

The model-based decomposition of income shocks, and its implied links with consumption changes, has also been used by Blundell *et al.* (2008) to study the evolution of income and consumption inequality using the PSID and the Consumer Expenditure Survey (CEX) for the period 1978–92. They document the consumption response to permanent and transitory income shocks across different education and age cohort groups; the estimated coefficients of the permanent shock vary between 0.4 and 0.95, while those of the transitory shock remain statistically indistinguishable from 0. They also conclude that the degree of insurance of the U.S. households against the two types of shocks remained unchanged over the sample period, while the relative *volatility* of the shocks has increased more for the insurable kind of income variation.

Applying a similar methodology, Jappelli and Pistaferri (2010b) use the Bank of Italy Survey of Household Income and Wealth (SHIW) for the period 1987–2006 to examine the benefits of the financial integration within the EMU in terms of potential improvements in the ability of households to smooth their consumption in the face of unanticipated income fluctuations. They find that household

consumption responses are not statistically different in the pre-EMU and post-EMU sub-samples, with the estimated consumption elasticity in the range of 0.7 to 1.0 for shocks to permanent income, and from 0 to 0.3 for shocks to transitory income.

We have found two consumption papers that use self-reported or *subjective assessments* of the persistence of household income, both applying methodologies other than that in the present study. Pistaferri (2001) uses data from the Italian SHIW for 1989–91 in which households reported their income in the survey year and their expectation of income in the following year. From this subjective information *and* additional assumptions on the development of household income across the life-cycle, income was decomposed into its temporary and persistent components. Pistaferri (2001) found no sensitivity to shocks with a transitory effect on income, while the estimated elasticity of shocks with a permanent effect is 0.57.[1]

Concurrently with the publication of the working paper version of this study (Kukk *et al.*, 2012), Sabelhaus and Ackerman (2012) published a study in which they derive the transitory income shock from self-reported income measures in the U.S. Survey of Consumer Finances (SCF). They investigate the extent to which the response to self-reported transitory income shocks contributed to the slowdown in the growth of food spending between 2004 and 2010. They estimate an Engel curve and find that a self-reported transitory income shock has a statistically significant effect on spending on food away from home. As the SCF makes it possible to derive the transitory income shock only for the relatively small fraction of households who report *unusually high* deviations of their current income from actual income, households with relatively small income deviations are left out of the study.

3. Dataset and Identification of Income Shocks

3.1. *The Estonian Household Budget Survey*

The Estonian HBS was conducted from 2002 to 2007 by Statistics Estonia, using a unified statistical methodology, which is outlined below and described in detail in ES (2012).[2] The rolling panel part of the Estonian HBS, used for empirical consumption modeling in this paper, consists of pairs of household observations on durable and non-durable consumption, different types of income, and various additional characteristics. The cross-sectional part of the Estonian HBS has previously been used by Kulikov *et al.* (2009) in an exploratory study of the saving behavior of Estonian households.

Each wave of the Estonian HBS comprises a representative cross-section of Estonian households based on regional stratification. The data are collected using a rolling panel structure: one half of the households are newly drawn from the

[1]Jappelli and Pistaferri (2000) also use subjective income expectations from the same survey, but test for excess sensitivity to *predicted* income innovations.

[2]The survey was discontinued in 2008–09 due to funding constraints and restarted in 2010 using a different sampling methodology. More information on the survey, including the list of official publications and methodology notes on both the pre-2008 and post-2009 samples, can be found on the website of Statistics Estonia (http://www.stat.ee/households).

population registry, while the second half is made up by re-interviewing the households that entered the survey the previous year. The data collection takes place in a sequence of 12 consecutive rounds, each corresponding to one calendar month and covering one-twelfth of the full annual sample. As a result of this survey design, the available time dimension of our panel is limited to two observations per household, spaced apart by exactly one calendar year. Due to the sample attrition and changes in response rates across different survey waves, the number of re-interviewed and newly drawn households displays some variations from year to year.

The panel contains observations on 2351 individual household units, each one observed at two time periods separated by one calendar year, for a total of 4,702 panel observations. This is the dataset resulting after some trimming of the original survey data. All households classified as self-employed have been excluded; this happens when the share of their business-related income in the total monthly income exceeds the threshold of 20 percent. The income of self-employed households is very volatile and subject to potentially large measurement errors.[3] Following the same argument a few observations with property income exceeding 20 percent of total income have also been excluded.[4] Finally, all observations where the head of the household is classified as inactive in the labor market in either of the time periods have been removed from the sample.[5]

The Estonian HBS contains a detailed breakdown of the *monthly* income and consumption figures of each household. The after-tax household income is composed of wage income, business-related income, property income, transfers, and other income sub-categories reported for the month of the interview. The total spending on household consumption is the sum of spending reported for the month of the interview on 12 different consumption expenditure groups; cf. the COICOP/HICP categories (Eurostat, 2014a).

We use two distinct monthly measures of after-tax household income which are available in the Estonian HBS: the *current* monthly income, and the *regular* income. The interviewed household is first asked to provide its current (or realized) income for the month of the interview. This information is based on detailed listings of different income sources to enhance the reliability of the data provided. The household is subsequently asked to provide an estimate of the regular income. The question is the following: "What is the usual amount of money at the disposal of your household during one month from all sources of income?" This question was asked *after* the household had stated its current monthly income.

The sequencing of the two questions should ensure that the household is aware of the differences between the two income concepts. This is confirmed by the fact that only 6 percent of the interviewed households report the same current and

[3]Kukk and Staehr (2014) refer to income under-reporting of self-employed households in the Estonian HBS. Krueger and Perri (2010) document substantial differences in the observed labor income volatility between self-employed and non-self-employed households in the PSID and SHIW datasets.

[4]This income category includes rents from owned land and real estate, interest income on deposits and investments, and intellectual property income. For the rest of the observations, property income comprises on average less than 0.2 percent of households' total income and therefore should not markedly affect the estimations.

[5]This partly addresses issues related to the possible non-separability of consumption and leisure in the utility of the representative household (see Attanasio and Weber, 2010).

regular income. In total 76 percent of the households report current income to be higher than regular income and 18 percent report current income to be lower than regular income. It cannot be ruled out that some households will equate regular income with the income from their main employment so that current income typically will be larger than regular income. The frequency of cases where current income is lower than regular income suggests, however, that this is not a very prevalent problem.

In any case, there is no means to ensure that parts of the different income components are correctly reported. The possibility of erroneous reporting compels us to examine the properties of the two income components in detail in the next subsection. It must be emphasized, however, that the purpose of this study is to investigate the consumption decision of households when the income shocks are derived from *subjective* measures of different income measures.

The two separate income measures make it possible to disentangle shocks with a long-term effect on income and shocks with a short-term effect based on the subjective assessment of the households. The methodology used for identifying household income shocks is discussed in Section 3.2. Although we are able to distinguish income shocks with different persistence, the survey does not provide any information on whether or not the income changes are expected. We discuss the implications for our results in the next subsection when the income shocks are computed.

On the consumption side, this paper takes advantage of two household consumption figures provided by the survey: full monthly household consumption, which covers all 12 COICOP/HICP sub-groups; and non-durable monthly household consumption, which is the sum of expenditures on food, non-alcoholic and alcoholic beverages, tobacco, clothing and footwear, housing (excluding regular maintenance and repair), transport services and fuel, newspapers, books and magazines, pet food, hotels, and eating out. Most of our results are based on the non-durable consumption measure, which is standard practice in the empirical consumption literature, but we also carry out robustness checks using the total monthly household consumption figure.[6]

In line with the theoretical model and the empirical literature, we convert all nominal income and consumption variables to real values. To this end, the monthly HICP price index for 2002–07 is used as a deflator (Eurostat 2014b; variable name: *prc_hicp_midx*, index 2005 = 1). We follow the convention in the empirical consumption literature and express the consumption and income variables in logarithms.

Table 1 shows summary statistics of the main consumption and income variables. Here and henceforth we use the subscripts in the following way: the index i refers to an individual household in the panel; the index $t \in \{1, 2\}$ refers to different observations of each individual household, where $t = 1$ denotes the first time the household is surveyed and $t = 2$ the second time the household is surveyed.

[6]The dynamics of durable consumption are in general different from those of non-durable consumption (cf. Bertola and Caballero, 1990). Since durable consumption goods deliver a stream of services lasting for many time periods, the correct way to account for them is to impute these service streams using an empirical procedure. We do not pursue this route, opting instead for a simpler way of separating the two kinds of consumption expenditures in our empirical models.

TABLE 1
MAIN VARIABLES IN THE DATASET, HOUSEHOLD-SPECIFIC DATA

Variable	Definition	Mean	S.D.
$\log C_{it}^{tot}$	Logarithm of real monthly total consumption expenditures	8.952	0.680
$\log C_{it}^{nd}$	Logarithm of real monthly non-durable consumption expenditures[a]	8.494	0.595
$\log Y_{it}$	Logarithm of real monthly after-tax income	9.151	0.636
$\log Y_{it}^{reg}$	Logarithm of real regular monthly after-tax income	9.018	0.558
$\log Y_{it}^{temp}$	Logarithm of real temporary monthly after-tax income, $\log Y_{it}^{temp} \equiv \log Y_{it} - \log Y_{it}^{reg}$	0.133	0.361

Notes: The real variables are in 2005 prices. In the sample period, the *kroon* (isocode EEK) was the currency in Estonia; the exchange rate was fixed at 15.65 EEK for 1 EUR.

[a]Expenditures on non-durable consumption include expenditures on food, alcohol, clothes and footwear, non-durable housing expenses, public transport and fuels, journals and magazines, pet food, eating out, travel and tourism expenses.

The calendar times of the interviews are not explicitly indexed, but a set of annual dummies is included in the empirical consumption models in Section 4 as a way of capturing business-cycle effects.

Beyond the variables directly stemming from the Estonian HBS, the last row of Table 1 also includes a measure of the *temporary income* of the household. The log temporary income, $\log Y_{it}^{temp}$, is defined as the difference between the household's log current income and its log regular income, that is, $\log Y_{it}^{temp} = \log Y_{it}^{cur} - \log Y_{it}^{reg}$.

The average real after-tax income Y_{it} amounts to 9424 EEK (602 EUR) per month, while the average real *regular* after-tax income Y_{it}^{reg} amounts to 8259 EEK (527 EUR) per month within the sample period. This means that the current income realized in the month on average exceeds the regular income by 14.1 percent.[7]

The discrepancy between the averages of the current and regular income measures is substantial, but it must be recalled that these measures are derived directly from the households' *subjective* assessments. The temporary income measure is equally a subjective measure computed without any constraints or statistical filtering. The fact that the realized income outpaced the regular income is likely the result of the very fast growth in real and nominal incomes during the sample period. Real GDP grew on average by 8.1 percent per year from 2002 to 2007 (Eurostat 2014b, code: *nama_gdp_k*). Nominal wages grew by 14.9 percent per year and real wages by 9.6 percent per year during the same period (Statistics Estonia 2014, code: *WS5311*; Eurostat 2014b, code: *prc_hicp_aind*).[8]

Apart from income and consumption data, the Estonian HBS contains a wealth of information about different household characteristics, including socio-demographic attributes (age of the head of the household and family size), variables for consumption characteristics (dummies for above-average or

[7]This also follows from the approximation afforded by the average of the log real *temporary* monthly after-tax income, which takes the value 0.133 or 13.3 percent.

[8]The fact that the average of the log real temporary income is substantially above 0 will not affect the results of the consumption estimations in Section 4 since the temporary income shocks entering the estimations are in all cases demeaned (cf. Section 3.2). The interpretation of the results may however depend on whether or not households correctly separate the temporary and persistent components.

below-average level) and indicators of economic affluence (participation in the labor market, property ownership flags, and liquidity position). A brief summary of these ancillary variables, which are mainly used as control variables in the empirical models in Section 4, is provided in Table A.1 in Appendix A.

3.2. *Identification of Income Shocks*

The feature of the Estonian HBS which sets it apart from most other micro-econometric datasets is the availability of two different income measures, which are assessed by the household as exhibiting different persistence. The empirical strategy used in Section 4 to estimate consumption sensitivities relies on the identification of shocks to the two income processes, cf. equations (1) and (2).

In our dataset the current income comprises two income measures which are available for computing the shocks: the regular income, $\log Y_{it}^{reg}$, which is assessed by the household as being relatively persistent, and the temporary income, $\log Y_{it}^{temp}$, which is assessed as exhibiting little persistence.

The shocks to the two income processes are computed in the following way:

$$\begin{aligned} \log Y_{it}^{reg} &= \alpha^{reg} + \rho^{reg} \log Y_{it-1}^{reg} + \nu_{it}^{reg} \\ \log Y_{it}^{temp} &= \alpha^{temp} + \rho^{temp} \log Y_{it-1}^{temp} + \nu_{it}^{temp}. \end{aligned} \tag{3}$$

The index i depicts the household, while the index t indicates whether the specific variable refers to the first or the second interview round ($t = 1$ or 2). The terms ρ^{reg} and ρ^{temp} are autoregressive coefficients and ν_{it}^{reg} and ν_{it}^{temp} are the shocks of the corresponding income processes. Depending on the autoregressive coefficients, the two income shocks are expected to have different impacts on the household consumption change between the two time periods; cf. equations (1) and (2).

The regular income variable conveys a perception of households' average income over a certain time span, possibly taking into account income expectations in the near future. The precise formulation of the question, however, leaves some open ends, as different respondents are likely to have different time horizons in mind when reporting their regular income. This makes it difficult to attain a clear picture of the persistence of the regular income measure implied by the survey responses.

At the same time, temporary income is defined as a residual of the current income left after subtraction of the regular income component; by construction it represents a highly idiosyncratic part of the household income. However, because this income measure is a linear combination of the current income $\log Y_{it}^{cur}$ and the regular income $\log Y_{it}^{reg}$, its persistence remains unknown: it is linked to the underlying persistence of the current and regular income variables.

In order to ascertain the empirical properties of regular and temporary household income, and to compute the shocks to the two processes, we estimate the coefficients in (3) using first micro data and then aggregated data. The results are shown in Table 2.

First, we employ a pooled OLS estimation to obtain estimates of the two persistence coefficients ρ^{reg} and ρ^{temp}. As noted, the time dimension of our panel is limited to two observations per household, making it impossible to control fully

TABLE 2

COEFFICIENT ESTIMATES OF PERSISTENCE OF REGULAR AND TEMPORARY INCOME

	(1) Regular Income	(2) Temporary Income	(3) Regular Income	(4) Temporary Income
$\hat{\rho}^{reg}$	0.815***	..	0.818***	..
	(0.015)		(0.072)	
$\hat{\alpha}^{reg}$	1.768***	..	1.645**	..
	(0.132)		(0.647)	
$\hat{\rho}^{temp}$	..	0.253***	..	–0.026
		(0.038)		(0.125)
$\hat{\alpha}^{temp}$	..	0.087***	..	0.139***
		(0.009)		(0.021)
Wald test (F-stat)	159.0	..	6.40	..
	[0.000]		[0.014]	
R^2	0.634	0.065	0.678	0.001
No. of obs.	2351	2351	71	71

Notes: OLS estimation of equation (3) with $\log Y_{it}^{reg}$ or $\log Y_{it}^{temp}$ as dependent variable. The coefficient $\hat{\rho}^{reg}$ is the estimated persistence for regular income and $\hat{\rho}^{temp}$ is the estimated persistence for temporary income, while $\hat{\alpha}^{reg}$ and $\hat{\alpha}^{temp}$ denote the respective constants. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, **, and * indicate that the coefficient is statistically different from 0 at the 1%, 5%, and 10% level, respectively. The Wald test shows the F-statistic and in square brackets below the p-value of the null hypothesis that the coefficient of the regular income is equal to one.

for time-invariant household heterogeneity by employing the fixed effects estimator. We make use of additional control variables to examine the issue in some detail.

Second, acknowledging the risk of biased results from the pooled OLS estimator, we depart from household-specific data and use aggregated data series. We average out household heterogeneity in our dataset by taking means of $\log Y_{it}^{reg}$ and $\log Y_{it}^{temp}$ across all households interviewed in each survey month from 2002 to 2007, obtaining two time series with 72 monthly observations each. An OLS estimation is then used to estimate the coefficients ρ^{reg} and ρ^{temp} from the averaged data.

Columns (1) and (2) in Table 2 show the estimated coefficients of the two income processes in (3) using the pooled OLS estimation. There are indeed substantial differences in the persistence of the two income variables: the estimate of ρ^{reg} in Column (1) is 0.81, while the estimate of ρ^{temp} in Column (2) is 0.25. Figure B.1 in Appendix B shows estimates of the two coefficients based on estimations undertaken for each month across the sample period. Although the confidence intervals of the rolling estimates are wider because the samples are smaller, on average they remain close to their full sample levels.

As mentioned previously, the results in Columns (1) and (2) of Table 2 are susceptible to unobserved household heterogeneity. While unable to take this heterogeneity into account fully with the fixed effects estimator, we examine the extent of the problem by adding control variables to our baseline regression in Table 2. The results are shown in Tables C.1 and C.2 in Appendix C. The estimate of ρ^{reg} is somewhat sensitive to the set of controls, but it remains in the interval of

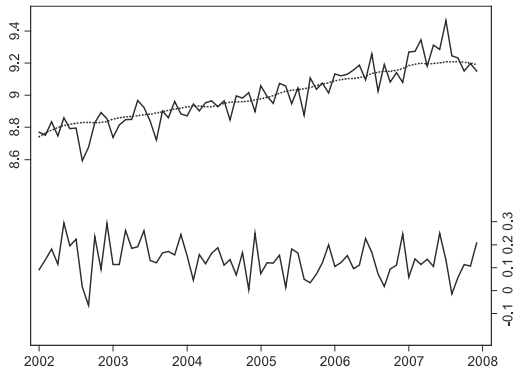


Figure 1. Variables $\log Y_{it}^{reg}$ (Left Scale) and $\log Y_{it}^{temp}$ (Right Scale) Averaged Across Households and the Real Gross Domestic Product (Dotted Line)

Notes: Monthly data. The real gross domestic product is from Eurostat (2014b, variable name: *namq_gdp_k*). Monthly data are obtained by interpolation from quarterly series using flexible polynomials, rescaled to match the average regular income level.

0.7–0.8 across various specifications in Table C.1. The estimate of ρ^{temp} remains broadly unchanged around the value of 0.25 for all model specifications in Table C.2.

We also implemented estimations for different sub-samples separately, for example for households with high and low education level of the household head, and for households with income below or above the median. The estimation results are provided in Tables C.3 and C.4 in Appendix C. The results remain very similar across the samples, lending support to our conjecture that the unobserved household heterogeneity does not have a substantial impact on the estimations of income persistence.

Columns (3) and (4) in Table 2 provide estimates of ρ^{reg} and ρ^{temp} obtained from the aggregated $\log Y_{it}^{reg}$ and $\log Y_{it}^{temp}$ series. The two series are displayed in Figure 1, showing the overall development of the two household income streams across the sample period from 2002 to 2007. The regular income appears to be closely following the trend of real gross domestic product, while the temporary income variable lacks any trend and exhibits no apparent dynamic structure. The estimate of ρ^{reg} in Column (3) confirms the persistence of the regular income stream obtained using the pooled data in Column (1). On the other hand, the temporary income persistence coefficient ρ^{temp} in Column (4) is now not statistically different from 0, suggesting that the corresponding pooled data result may be biased due to the omitted household heterogeneity term in the presence of a lagged dependent variable.[9]

[9]The autoregressive coefficients in columns (3) and (4) of Table 2 are likely to be underestimated due to a well-known downward small-sample bias of the least-squares estimator in the linear dynamic models with lags (see Shaman and Stine, 1988).

In summary, our results in Table 2 suggest that the regular income process is quite persistent, with the coefficient ρ^{reg} lying in the interval 0.7–0.8. By contrast, the temporary income process exhibits little or no persistence, with ρ^{temp} falling in the interval of 0–0.25. The estimated range of ρ^{temp} implies that the shock to temporary income, v_{it}^{temp}, is likely to have little effect on the household's consumption plans.

The lack of household-specific estimates for the persistence of regular and temporary income streams calls for a judicious approach to the estimation of the corresponding income shocks. Therefore, we consider several different assumptions regarding the persistence of the two income streams. For each assumption, we compute residuals from equation (3) and use these shocks when estimating the consumption responses.

The baseline assumption is that $\rho^{reg} = 1$ and $\rho^{temp} = 0$. Under this assumption, the shock to regular income, v_{it}^{reg}, has a permanent effect on income, that is, $\log Y_{it}^{reg}$ is a unit root process, while the shock to temporary income, v_{it}^{temp}, affects income for just a single time period, that is, $\log Y_{it}^{temp}$ is a white noise process. This choice of baseline is also a direct reflection of the households' own assessment of the persistence of the two income components.

We also consider three alternative assumptions regarding the persistence of the income processes, cf. the estimations in Table 2. The first alternative assumes that $\rho^{reg} = 0.9$ and $\rho^{temp} = 0$; the second alternative that $\rho^{reg} = 0.8$ and $\rho^{temp} = 0$; and the third alternative that $\rho^{reg} = 0.8$ and $\rho^{temp} = 0.25$. These sets of assumptions broadly cover the different estimates of the two persistence coefficients. They are used to examine the impact on the estimated consumption responses of different assumptions regarding the persistence of the income shocks.

Table 3 provides the statistics for the shocks v_{it}^{reg} and v_{it}^{temp}, which are computed using the four alternative assumptions about the persistence of the shocks. The means of the empirical income shocks are 0 in all cases due to the way the shocks are computed. The standard deviations are very similar across the different persistence assumptions in Columns (1)–(4), but are substantially higher than those found in most other studies; see the overview of Meghir and Pistaferri (2011). The high standard deviations may reflect the rapid changes in household income during the economic boom as discussed in Section 3.1. This might lead to higher variability in both income shocks.

TABLE 3

Descriptive Statistics of Shocks to Regular and Temporary Income, Different Assumptions about Their Persistence

	(1) $\rho^{reg} = 1$, $\rho^{temp} = 0$		(2) $\rho^{reg} = 0.9$, $\rho^{temp} = 0$		(3) $\rho^{reg} = 0.8$, $\rho^{temp} = 0$		(4) $\rho^{reg} = 0.8$, $\rho^{temp} = 0.25$	
	v_{it}^{reg}	v_{it}^{temp}	v_{it}^{reg}	v_{it}^{temp}	v_{it}^{reg}	v_{it}^{temp}	v_{it}^{reg}	v_{it}^{temp}
S.D.	0.355	0.360	0.343	0.360	0.340	0.360	0.340	0.348
Correlation	−0.185		−0.188		−0.186		−0.240	

Notes: The empirical income shocks v_{it}^{reg} and v_{it}^{temp} are computed by inserting the persistence assumptions in the column headings into equation (3).

It is notable that the standard deviations of temporary and regular shocks are of the same magnitude, which may be at odds with the understanding that more persistent income shocks would have smaller variances. The reason may again be related to the rapid increase in average incomes during the sample period which are likely to have led to substantial changes in the assessment of regular income levels. In addition, as we are using survey data the large standard deviations of the derived shocks may also indicate the presence of the measurement errors.

The shocks to regular and temporary income are negatively, but weakly, correlated. A negative correlation is also found in Hryshko (2013), who estimates the correlation coefficient between permanent and transitory shocks to be –0.5 in data from the SHIW.

The theoretical model of intertemporal consumption choice in equation (1) assumes that the shocks to the income processes are unanticipated. There is no information in the survey on this issue, but the empirical equivalents v_{it}^{reg} and v_{it}^{temp} were constructed so as to make it probable that this assumption also holds in the empirical model. Furthermore, we investigated the descriptive power of the socio-demographic and economic variables for the shocks. The adjusted R^2 is below 0.02 for both income shocks and the value of the F-statistic is below 5, indicating a very modest explanatory power of the variables (not shown). This suggests that the derived shocks are to a large extent unexplainable by observed variables.

4. Consumption Estimations

4.1. *Estimations of the Standard Consumption Model*

The empirical model originates from equation (2) where control variables are added in the form of preference shifters and time fixed effects (to capture aggregate economic developments). The following equation is used for estimating consumption sensitivities to shocks to different income components:

$$\Delta \log C_{it} = Z_{it}'\gamma + \beta_1 v_{it}^{reg} + \beta_2 v_{it}^{temp} + \varepsilon_{it}. \tag{4}$$

The dependent variable $\Delta \log C_{it}$ is the change in the logarithm of consumption between two time periods. The dependent variable is taken to be non-durable consumption in most cases, but it is total consumption in some robustness analyses. The control variables in the column vector Z_{it} are annual time dummies and preference shifters in the form of the change in household size and the logarithm of the age of the head of the household (cf. Attanasio, 1999). The annual time dummies seek to capture aggregate effects on household consumption.[10] (Given the inclusion of time dummies for each period, no constant is included.) The vector γ contains the coefficients of the control variables. The coefficient of the shock to regular income is β_1 and the coefficient of the shock to temporary income is β_2. Finally, ε_{it} is an error term.

[10]We also tested the inclusion of monthly dummies for seasonal effects, but found that the dummies were typically statistically insignificant, while the other estimation results remained essentially unchanged. The reason for the absence of seasonality effects is likely to be that the regressions include changes in consumption and income from the *same month* the previous year.

TABLE 4

RESPONSE OF NON-DURABLE CONSUMPTION TO INCOME SHOCKS, DIFFERENT ASSUMPTIONS REGARDING THE PERSISTENCE OF THE INCOME PROCESSES

	(1) $\rho^{reg} = 1$ $\rho^{temp} = 0$	(2) $\rho^{reg} = 0.9$ $\rho^{temp} = 0$	(3) $\rho^{reg} = 0.8$ $\rho^{temp} = 0$	(4) $\rho^{reg} = 0.8$ $\rho^{temp} = 0.25$
v_{it}^{reg}	0.260***	0.259***	0.243***	0.266***
	(0.033)	(0.034)	(0.033)	(0.034)
v_{it}^{temp}	0.121***	0.120***	0.116***	0.176***
	(0.029)	(0.028)	(0.028)	(0.034)
Wald test (F-stat)	11.81	11.61	9.86	4.59
	[0.001]	[0.001]	[0.002]	[0.032]
No. of obs.	2351	2351	2351	2351
R^2	0.076	0.073	0.069	0.077

Notes: OLS estimation of equation (4) with $\Delta \log C_{it}$ as dependent variable. The variable v_{it}^{reg} is the regular income shock and v_{it}^{temp} is the temporary income shock derived from equation (3). Household size, log of age of household head, and year dummies are included in the estimations but are not shown in the table. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, **, and * indicate that the coefficient is statistically different from 0 at the 1%, 5%, and 10% level, respectively. The Wald test shows the F-statistic and in square brackets below the p-value of the null hypothesis that the coefficients of the shocks to regular and temporary income are identical.

Table 4 shows the first set of estimation results in which changes in non-durable consumption are explained by shocks to income of different persistence and the set of control variables. Column (1) provides the results of the baseline estimation in which the shocks to regular income are assumed to be fully persistent ($\rho^{reg} = 1$) and the shocks to temporary income to be without persistence ($\rho^{temp} = 0$). A shock to regular income of 10 percent induces a consumption increase of 2.6 percent in the same period. The results are lower than those obtained in the studies that use econometric identification of the shocks, listed in Section 2. According to these studies, the elasticity of consumption to a permanent income shock is between 0.6 and 1. In our study, the self-reported regular income is not fully persistent and the estimated coefficient should therefore be lower.

It also follows from Column (1) in Table 4 that a shock to temporary income of 10 percent leads to a 1.2 percent increase in consumption. Although studies that use statistical decomposition typically report a negligible consumption response, studies based on quasi-experiments find a consumption response to transitory income of a magnitude comparable to our estimate, cf. the comprehensive review by Jappelli and Pistaferri (2010a).

Nevertheless, among households in Estonia, non-durable consumption is more sensitive to a shock to the highly persistent income component than to a shock to the less persistent income component. The difference between the two coefficients is statistically significant at the 1 percent level. The results reflect the fact that households distinguish between shocks to income processes that contain different levels of persistence.

Columns (2)–(3) in Table 4 show the estimation results when different assumptions about the persistence of the regular income process are used. The results are qualitatively similar to those of the baseline estimation, indicating that

the results are rather insensitive to the persistence of the permanent income variable. Column (4) shows the results when temporary income is assumed to exhibit some persistence. Overall the results in Columns (2)–(4) suggest that an inability to identify the true income process of the households does not affect the results of the consumption estimations. The results in Table 4 are robust to a large number of specification changes and to the inclusion of different control variables.[11]

The estimated coefficient of a shock to temporary income (0.12) is comparable to the results in Jappelli and Pistaferri (2010b), based on econometric decomposition of different income processes, as they estimated the coefficient to be in the range of 0 to 0.3. Blundell *et al.* (2008) and Pistaferri (2001) found no response of consumption to a shock with transitory effect on income. The estimated coefficient of a shock to regular income (0.26) is smaller than that estimated in studies where the coefficient for the full sample is estimated to be around 0.6. As the latter are estimated for shocks with a permanent effect on income, while the regular income used in this paper exhibits persistence that is lower than 1, the estimation results should be different. The consumption response is expected to decrease fast when the persistence coefficient falls below 1 (cf. Kaplan and Violante, 2010; Hryshko, 2013).

As discussed in Section 3.1, the decomposition of income into regular and temporary components relies on the households' own assessments, and we can therefore not rule out the possibility of measurement errors due to the households mixing up the two income components. If the households include part of their temporary income in the regular income measure, then we would expect the consumption response to the derived regular income shock to be relatively low. If, on the other hand, households do not include all of their persistent income in the regular income measure, we would expect a positive consumption response to a temporary income shock. The sequencing of the income questions should, however, diminish the risk that the households mix up the two income concepts when being interviewed. Moreover, possible measurement errors are arguably of less concern in the present analysis since the aim is to uncover the consumption response to income shocks derived from the subjective income measures of households.

Moreover, if the income shocks contain an anticipated component, the consumption response is expected to be lower as households have already adjusted to the anticipated income changes (see discussion in Jappelli and Pistaferri, 2010a). On top of that, if households are insured against some of their income shocks, the consumption response to regular shocks is smoothed, as noted by Blundell *et al.* (2008).

In conclusion, given the lower persistence of the regular income than in many other studies, potential measurement errors in the income components, possible anticipated income changes, and partial insurance, the estimation results in this paper using self-reported income measures are comparable with those in studies that use statistical decomposition.

[11]The robustness tests are available from the authors upon request.

4.2. *The Response of Consumption to Lagged Shock to Temporary Income*

The consumption response to a temporary income shock found in Section 4.1 cannot be explained by the presence of persistence in the temporary income process. Even if the persistence is 0.25, the upper limit in Table 2, the shock still dies out fast and has very little effect on lifetime earnings. Moreover, although it cannot be completely ruled out that some of the temporary income shock is anticipated, this cannot explain the significant response of consumption. An anticipated temporary income shock should, in line with an unanticipated temporary income shock, have a negligible impact on consumption (Jappelli and Pistaferri, 2010a).

One possibility is that households use short time horizons in their consumption decisions, that is, they divide the windfall income between a limited number of future time periods. The shorter than life-time horizon is also mentioned by Friedman (1957) and can still be considered to be forward-looking behavior. Another possibility is that households react to a shock to temporary income only in the time period in which the shock occurs, that is, they do not fully smooth consumption in this case. This type of behavior is considered to be either myopic or induced by some constraints on forward-looking behavior such as liquidity constraints (Attanasio and Weber, 2010).

The consumption specification in equation (4) does not allow us to shed light on the reasons for the consumption response to the temporary income shock. In an attempt to shed further light on the issue, we include the lagged temporary income shock. In case the consumption response to temporary income shocks is due to a short time horizon of households, the coefficient of the lagged temporary income variable is expected to be statistically insignificant as the temporary income is divided proportionally across time periods. If the consumption follows current income, the sign of the significant coefficient is expected to be negative.

The consumption specification when the lagged temporary income shock v_{it-1}^{temp} is included takes the form:

$$\Delta \log C_{it} = Z_{it}\gamma + \beta_1 v_{it}^{reg} + \beta_2 v_{it}^{temp} + \beta_3 v_{it-1}^{temp} + \varepsilon_{it}. \tag{5}$$

As there are only two observations per household, we have to restrict the persistence of the shock to temporary income to 0 in order to include the lagged temporary income variable.

Table 5 shows the results of estimations of equation (5). The estimated coefficient of the lagged shock to temporary income is negative and statistically significant. Moreover, inclusion of the lagged shock changes the estimated coefficients for current regular and temporary shocks.

The coefficient of the lagged shock to temporary income is negative and, in numerical terms, close to the coefficient of the current shock to temporary income. This holds irrespective of the assumption of persistence in the shock to regular income. The result suggests that consumption reacts positively to a shock to positive temporary income only in the same period when the shock appears. In the following period the consumption has to retreat by the same magnitude as it responded to the shock during the previous period.

TABLE 5

ESTIMATIONS WHEN LAGGED SHOCK TO TEMPORARY INCOME IS INCLUDED

	(1) $\rho^{reg} = 1$ $\rho^{temp} = 0$	(2) $\rho^{reg} = 0.9$ $\rho^{temp} = 0$	(3) $\rho^{reg} = 0.8$ $\rho^{temp} = 0$
v_{it}^{reg}	0.329***	0.328***	0.308***
	(0.034)	(0.035)	(0.034)
v_{it}^{temp}	0.196***	0.194***	0.186***
	(0.036)	(0.035)	(0.035)
v_{it-1}^{temp}	−0.247***	−0.242***	−0.233***
	(0.030)	(0.030)	(0.030)
Wald test (F-stat)	9.96	11.11	8.38
	[0.002]	[0.002]	[0.004]
R^2	0.106	0.102	0.096
No. of obs.	2351	2351	2351

Notes: OLS estimation of equation (5) with $\Delta \log C_{it}$ as dependent variable. The variable v_{it}^{reg} is the regular income shock, v_{it}^{temp} is temporary income shock and v_{it-1}^{temp} is lagged temporary income shock derived from equation (3). Household size, log of age of household head, and year dummies are included in the estimations but not shown in the table. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, **, and * indicate that the coefficient is statistically different from 0 at the 1%, 5%, and 10% level, respectively. The Wald test shows the F-statistic and in square brackets below the p-value of the null hypothesis that the coefficients of the shocks to regular and temporary income are identical.

When the effect of a lagged shock to temporary income is taken into account, the coefficients of the current period income shocks increase relative to the results in Table 4. A shock that increases regular income by 10 percent increases consumption by 3.3 percent in the same month. A current period shock to temporary income of the same magnitude increases consumption by 2 percent. The coefficients of the regular and temporary income shocks of the same period are still significantly different, suggesting that households do indeed distinguish between income shocks to different income components and they incorporate this information into their consumption decisions.

The findings in Table 5 are robust to different assumptions about the identification of the income shocks, as illustrated by the estimations in Columns (2) and (3). Estimations with total consumption instead of non-durable consumption also produce qualitatively similar results; cf. Table D.1 in Appendix D. The response of total consumption to income shocks is larger than that of non-durable consumption and the lagged shock to temporary income is significant with a negative sign.

The results are robust to the inclusion of additional control variables that capture changes in the economic situation and the behavior of the household; see Table D.2 in Appendix D. These robustness tests are important as the dataset has only two observation points per household, which implies that the differenced specification in (5) has only one observation per household, ruling out any fixed effects estimation. The inclusion of variables capturing peculiarities in consumption, changes in employment status, household wealth or liquidity constraints do

not change the coefficients of the income shocks in substantial ways. The coefficients of consumption response are not affected by omitted variables bias with the notable exception of the lagged shock to temporary income.

The statistical and economic significance of the lagged shock to temporary income might relate to numerous explanations, such as myopic behavior of households or partial consumption insurance, for instance due to liquidity constraints. Previous studies have found evidence for both explanations (see the overview of Attanasio and Weber, 2010).

In standard models of intertemporal consumption choice, households react to the current and lagged negative transitory shock if they are liquidity constrained. As noted by Jappelli and Pistaferri (2010a), constrained households are not able to smooth declines in transitory income through borrowing.

In order to understand the impact of liquidity constraints on our results, we estimated equation (5) on two different sub-samples. We split the households into two sub-samples based on their *liquidity score*. The liquidity score is derived from subsets of questions about the ability of the household to finance consumption expenditures of various nominal values instantly by borrowing or by drawing down savings.[12] The score ranges from 0 to 6 and larger values indicate easier access to liquidity. We pooled the scores of 0–2, which correspond to tight constraints on consumption, as these households are not able to finance consumption expenditure of 15,000 EEK (959 EUR) instantly from either savings or credit. The second sub-sample consists of households with liquidity scores between 3 and 6, which indicate that a sufficient amount of savings or credit is available to finance consumption expenditures of values above 15,000 EEK. We label the sub-samples as households with low liquidity and high liquidity.

Table 6 presents the estimation results for the two sub-samples. It shows a higher consumption response to a transitory income shock among households with low liquidity, although the difference is not significant in statistical terms. Households with low liquidity exhibit a significantly higher response to an income shock to regular income than do households with high liquidity; the point estimates are 0.45 and 0.26, respectively. In the framework of partial insurance models, these results can be explained by constraints on self-insurance; consequently, households are not able to smooth their regular income. There might be other explanations for the results, such as differences in the persistence of regular income or the presence of anticipated income shocks, but a deeper analysis is beyond the scope of this paper. The overall result is that liquidity constraints do not explain the consumption response to the temporary shock and other explanations, including myopic behavior of households, may therefore be considered.

The results in this subsection are important for several reasons. First, most studies investigating the response of consumption to transitory income use only current variables, as lagged variables are seldom available. Our results show that omitting the lagged temporary income may bias the estimated coefficients of current income shocks downwards. This may be important when the estimated

[12]The survey asked about the ability to finance instant consumption expenditures for three values—1000 EEK (64 EUR), 5000 EEK (320 EUR), and 15,000 EEK (959 EUR)—and specifies the source of financing.

TABLE 6

SENSITIVITY OF NON-DURABLE CONSUMPTION TO SHOCKS BY DIFFERENT SUBSAMPLES

	(1) Low Liquidity	(2) High Liquidity
v_{it}^{reg}	0.451***	0.261***
	(0.063)	(0.040)
v_{it}^{temp}	0.222***	0.183***
	(0.052)	(0.045)
v_{it-1}^{temp}	−0.270***	−0.222***
	(0.042)	(0.043)
Wald test (F-stat)	8.17	2.83
	[0.004]	[0.093]
R^2	0.122	0.106
No. of obs.	863	1488

Notes: OLS estimation of equation (5) with $\Delta \log C_{it}$ as dependent variable. The variable v_{it}^{reg} is the regular income shock, v_{it}^{temp} is temporary income shock and v_{it-1}^{temp} is lagged temporary income shock derived from equation (3). Household size, log of age of household head, and year dummies are included in the estimations but are not shown in the table. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, **, and * indicate that the coefficient is statistically different from 0 at the 1%, 5%, and 10% level, respectively. The Wald test tests the null hypothesis that the difference between the coefficients of shocks to regular and temporary income is not statistically significant. The values in the square brackets are p-values.

coefficients of the standard intertemporal consumption model without a lagged temporary income variable are interpreted.

Second, the significant negative coefficient of the lagged temporary income shock in the sub-sample of households without liquidity constraints is not consistent with standard consumption models of forward-looking unconstrained households. Our results show that households tend to react to contemporaneous temporary income shocks during the same time period. The study of Hall and Mishkin (1982) is among the few that estimate the responses to both current and lagged transitory shocks. They obtain a negative sign for the lagged variable and explain it by alluding to the rule-of-thumb consumption.

5. CONCLUSION

A detailed understanding of the determinants of household consumption is important for an analysis of microeconomic welfare and for macroeconomic policy prescriptions. This paper analyzes how household consumption in Estonia reacts to different income shocks.

The main innovation of the analysis is the identification of income shocks of different persistence using information provided by the individual households. The identification of income persistence based on self-reported income measures is uniquely made possible by data in the Estonian HBS. This eliminates the need for

restrictive statistical decomposition, which typically utilizes additional assumptions about the co-movements of consumption and income.

The data sample consists of 2351 households interviewed twice during the period 2002–07, a period of rapid economic growth and increasing household income in Estonia. The two income measures are regular income and temporary income, where the latter is defined as the difference between the current and regular incomes reported by the households. In order to assess the properties of the income measures, different empirical methods are used to produce approximate estimates of their persistence. The analyses show that the households diligently assess the persistence of the two income measures. Shocks or innovations to the two different income processes can be calculated.

The starting point for the consumption estimations is a model with a shock to regular and temporary income. The consumption response to a shock affecting regular income is significantly higher than the response to a shock affecting temporary income. This result entails a different degree of consumption smoothing depending on the persistence of the income shocks, and this is consistent with standard models of intertemporal consumption choice. The estimations are robust to different consumption measures, to different degrees of persistence of the income processes, and to the inclusion of additional control variables.

Further estimations reveal, however, that consumption also reacts to the lagged shock in temporary income, which is not consistent with standard models of forward-looking households. The coefficient of the lagged shock to temporary income is negative and, in numerical terms, close to the coefficient of the contemporaneous shock to temporary income. Exploratory investigations rule out liquidity constraints as an important factor for the results, suggesting that other explanations such as myopic behavior should be considered. The upshot is in any case that the consumption choice is in part affected by factors that cannot be ascribed to intertemporal smoothing of consumption.

The identification of income shocks based on self-reported income measures is novel. This study contributes to the empirical literature on consumption behavior by estimating consumption responses to income shocks of different persistence given the subjective assessment of households. In spite of the unique shock identification and sample, the results are broadly in line with findings in earlier studies, namely that households react differently to shocks to income of different persistence. Nevertheless, as households respond to transitory income changes, the study presents new evidence that consumption also adjusts to the transitory income of the previous period. The finding helps to improve the understanding about the co-movement of consumption and income.

An important direction for future research is the estimation of different consumption responses to positive and negative shocks to income. Some empirical evidence suggests that households perceive changes in their financial situation asymmetrically, as found by Mastrogiacomo (2010) among others. Tests of asymmetry for consumption response can provide additional information on the extent of forward-looking behavior, and also on reasons for deviations from the standard model of intertemporal consumption choice. The challenges of implementing such tests are considerable and can be left for future research.

References

Aguila, E., O. Attanasio, and C. Meghir, "Changes in Consumption at Retirement: Evidence from Panel Data," *Review of Economics and Statistics*, 93, 1094–9, 2011.

Attanasio, O., "Consumption," in J. Taylor and M. Woodford (eds), *Handbook of Macroeconomics*, Elsevier, Rotterdam, 741–812, 1999.

Attanasio, O. and G. Weber, "Consumption and Saving: Models of Intertemporal Allocation and their Implications for Public Policy," *Journal of Economic Literature*, 48, 693–751, 2010.

Bertola, G. and R. J. Caballero, "Kinked Adjustment Costs and Aggregate Dynamics," in O. J. Blanchard and S. Fisher (eds), *NBER Macroeconomics Annual*, MIT Press, Cambridge, MA, 1990.

Blundell, R., L. Pistaferri, and I. P. Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, 98, 1887–921, 2008.

Brixiova, Z., L. Vartia, and A. Worgotter, "Capital Flows and the Boom-Bust Cycle: The Case of Estonia," *Economic Systems*, 34, 55–72, 2010.

Chao, H.-K., "Milton Friedman and the Emergence of the Permanent Income Hypothesis," *History of Political Economy*, 35, 77–104, 2003.

Deaton, A. S., "Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?" in T. F. Bewley (ed.), *Advances in Econometrics: Fifth World Congress*, Vol. 2, University Press, Cambridge, New York, and Melbourne, 121–48, 1987.

ES, "Household Budget Survey 2010: Methodology," Statistics Estonia Publication (http://www.stat.ee/dokumendid/59873), 2012.

Eurostat, "Statistics Explained. Statistical Glossary" (http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:COICOP/HICP), 2014a.

———, Statistics Database (http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home), 2014b.

Friedman, M., *A Theory of the Consumption Function*, Princeton University Press, Princeton and Oxford, 1957.

Gustavsson, M., "A New Picture of Swedish Earnings Inequality: Persistent and Transitory Components, 1960–1990," *Review of Income and Wealth*, 54, 324–49, 2008.

Hall, R. E. and F. Mishkin, "The Sensitivity of Consumption to Transitory Income: Evidence from PSID Households," *Econometrica*, 50, 461–81, 1982.

Heathcote J., K. Storesletten, and G. L. Violante, "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, 1, 319–54, 2009.

Hryshko, D., "Excess Smoothness of Consumption, Consumption Insurance, and Correlated Income Shocks," Mimeo, University of Alberta, 2013.

Hryshko, D., J. Luengo-Prado, and B. E. Sørensen, "House Prices and Risk Sharing," *Journal of Monetary Economics*, 57, 975–87, 2010.

Jappelli, T. and L. Pistaferri, "Using Subjective Income Expectations to Test for Excess Sensitivity of Consumption to Predicted Income Growth," *European Economic Review*, 44, 337–58, 2000.

———, "The Consumption Response to Income Changes," *Annual Review of Economics*, 2, 479–506, 2010a.

———, "Financial Integration and Consumption Smoothing," *Economic Journal*, 121, 678–706, 2010b.

Kaplan, G. and G. L. Violante, "How Much Consumption Insurance Beyond Self-Insurance?" *American Economic Journal: Macroeconomics*, 2, 53–87, 2010.

Krueger, D. and F. Perri, "How does Household Consumption Respond to Income Shocks?" Mimeo, University of Pennsylvania (http://economics.sas.upenn.edu/~dkrueger/research/KPItaly.pdf), 2010.

Kukk, M. and K. Staehr, "Income Underreporting by Households with Business Income. Evidence from Estonia," *Post-Communist Economies*, 26, 257–76, 2014.

Kukk, M., D. Kulikov, and K. Staehr, "Consumption Sensitivities in Estonia: Income Shocks of Different Persistence," Working Papers of Eesti Pank, No. 3/2012, 2012.

Kulikov, D., A. Paabut, and K. Staehr, "A Microeconometric Analysis of Household Saving in Estonia: Income, Wealth and Financial Exposure," in D. Mayes (ed.), *Microfoundations of Economic Success: Lessons from Estonia*, Edward Elgar, Cheltenham and Northampton, 190–241, 2009.

Mastrogiacomo, M., "Testing Consumers' Asymmetric Perception of Changes in Household Financial Situation," *Review of Income and Wealth*, 56, 327–50, 2010.

Meghir, C. and L. Pistaferri, "Earnings, Consumption, and Life Cycle Choices," in D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 4B, 773–854, Elsevier, Amsterdam, 2011.

Pistaferri, L., "Superior Information, Income Shocks, and the Permanent Income Hypothesis," *Review of Economics and Statistics*, 83, 465–76, 2001.

Quah, D., "Permanent and Transitory Movements in Labor Income: An Explanation for 'Excess Smoothness' in Consumption," *Journal of Political Economy*, 98, 449–75, 1990.
Sabelhaus, J. and S. Ackerman, "The Effect of Self-Reported Transitory Income Shocks on Household Spending," FEDS Working Paper No. 2012-64, Federal Reserve Board, 2012.
Sahm, C. R., M. D. Shapiro, and J. Slemrod, "Household Response to the 2008 Tax Rebate: Survey Evidence and Aggregate Implications," *Tax Policy and the Economy*, 24, 69–110, 2010.
Shaman, P. and R. Stine, "The Bias of Autoregressive Coefficient Estimators," *Journal of the American Statistical Association*, 83, 842–48, 1988.
Statistics Estonia, "Statistical Database: Economy," (http://pub.stat.ee/px-web.2001/I_Databas/Economy/databasetree.asp), 2014.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix A: Variable definitions
Table A.1: Definitions of additional variables used in the empirical models of household consumption behaviour
Appendix B: Identification of income shocks in the Estonian HBS
Figure B.1: Rolling monthly estimates of ρ^{reg} (left scale) and ρ^{temp} (right scale) together with their 95 percent confidence intervals (dotted line) and estimates using the full sample (dashed lines)
Appendix C: Robustness checks of income persistence estimations
Table C.1: Regular income persistence coefficient with control variables
Table C.2: Temporary income persistence coefficient with controls
Table C.3: Income persistence coefficients for different sub-samples
Table C.4: Income persistence coefficients for different sub-samples
Appendix D: Robustness tests of estimations with lagged shock to temporary income
Table D.1: Sensitivity of total consumption to shocks to income processes of different persistence
Table D.2: Robustness test of the baseline regression to different sets of control variables

On-line Appendix

Estimating consumption responses to income shocks of different persistence using self-reported income measures

Review of Income and Wealth

Merike Kukk, Dmitry Kulikov and Karsten Staehr

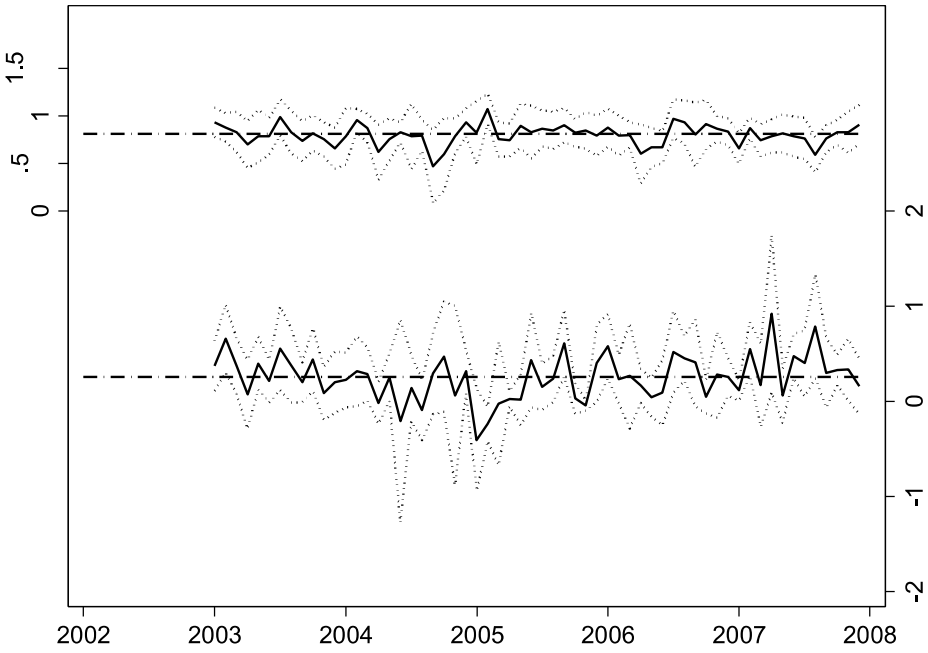
Appendix A. Variable definitions

Table A.1. Definitions of additional variables used in the empirical models of household consumption behaviour

Variable	Definition
$lage_{it}$	Logarithm of the age of the household head
$hhsize_{it}$	Number of household members
$belowc_{it}$	Dummy = 1 if household's consumption in a given month is lower than the regular level, otherwise = 0
$abovec_{it}$	Dummy = 1 if household's consumption in a given month is higher than the regular level, otherwise = 0
$partempl_{it}$	Dummy = 1 if not all adult household members are currently employed, otherwise = 0
$renting_{it}$	Dummy = 1 if the household lives in a rented dwelling, otherwise = 0
$realest_{it}$	Dummy = 1 if the household owns real estate in addition to its primary residence, otherwise = 0
$liquid_{it}$	Household liquidity score from 0 to 6. This variable is derived from answers on the ability of the household to finance consumption expenditures instantly of different nominal values by getting loans or drawing on own funds; cf. Kulikov et al. (2009). Larger values indicate easier access to liquidity

Appendix B. Identification of income shocks in the Estonian HBS

Figure B.1. Rolling monthly estimates of ρ^{reg} (left scale) and ρ^{temp} (right scale) together with their 95% confidence intervals (dotted line) and estimates using the full sample (dashed lines)



Appendix C. Robustness checks of income persistence estimations

Table C.1 Regular income persistence coefficient with control variables

	(1)	(2)	(3)	(4)	(5)	(6)
$\log Y_{it-1}^{reg}$	0.815***	0.770***	0.769***	0.767***	0.761***	0.681***
	(0.015)	(0.017)	(0.018)	(0.018)	(0.018)	(0.023)
$lage_{it-1}$	..	-0.041	-0.041	-0.041	-0.060**	-0.073***
		(0.027)	(0.027)	(0.027)	(0.028)	(0.028)
$hhsize_{it-1}$	..	0.046***	0.046***	0.045***	0.043***	0.059***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
$partempl_{it-1}$	..	..	-0.008	-0.008	-0.010	-0.002
			(0.016)	(0.016)	(0.016)	(0.016)
$abovec_{it-1}$	..	..	..	0.038***	0.039***	0.035**
				(0.015)	(0.015)	(0.015)
$belowc_{it-1}$	..	..	..	0.036	0.035	0.036
				(0.028)	(0.028)	(0.028)
$realest_{it-1}$	..	..	..	..	0.039*	0.031
					(0.021)	(0.020)
$renting_{it-1}$	..	..	..	..	-0.081***	-0.074***
					(0.027)	(0.027)
$liquid_{it-1}$	..	..	..	..	..	0.040***
						(0.006)
Constant	1.768***	2.197***	2.209***	2.216***	2.349***	2.949***
	(0.132)	(0.180)	(0.184)	(0.184)	(0.194)	(0.228)
R^2	0.634	0.642	0.642	0.643	0.644	0.652
No. of obs.	2351	2351	2351	2351	2351	2351

Notes: OLS estimation of eq. (3) with $\log g_{it}^{reg}$ as dependent variable. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the corresponding coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

Table C.2 Temporary income persistence coefficient with controls

	(1)	(2)	(3)	(4)	(5)	(6)
$\log Y_{it-1}^{temp}$	0.253***	0.250***	0.250***	0.240***	0.240***	0.240***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
$lage_{it-1}$	..	-0.038*	-0.038*	-0.040*	-0.029	-0.028
		(0.023)	(0.023)	(0.023)	(0.024)	(0.024)
$hhsize_{it-1}$	..	0.006	0.006	0.005	0.007	0.007
		(0.007)	(0.007)	(0.006)	(0.007)	(0.007)
$partempl_{it-1}$	..	..	-0.007	-0.008	-0.007	-0.008
			(0.015)	(0.015)	(0.015)	(0.015)
$abovec_{it-1}$	..	..	..	0.043***	0.044***	0.044***
				(0.016)	(0.016)	(0.016)
$belowc_{it-1}$	..	..	..	-0.021	-0.019	-0.019
				(0.024)	(0.024)	(0.024)
$realest_{it-1}$	..	..	..	..	0.013	0.014
					(0.020)	(0.020)
$renting_{it-1}$	..	..	..	..	0.071**	0.070**
					(0.035)	(0.035)
$liquid_{it-1}$	..	..	..	..	..	-0.001
						(0.004)
Constant	0.087***	0.215**	0.217**	0.215**	0.162*	0.164*
	(0.009)	(0.092)	(0.092)	(0.092)	(0.098)	(0.099)
R^2	0.065	0.066	0.066	0.07	0.072	0.072
No. of obs.	2351	2351	2351	2351	2351	2351

Notes: OLS estimation of eq. (3) with $\log Y_{it}^{temp}$ as dependent variable. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the corresponding coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

Table C.3. Income persistence coefficients for different sub-samples

	(1)	(2)	(3)	(4)
	Regular income	**Regular income**	**Temporary income**	**Temporary income**
	Low education level	High education level	Low education level	High education level
$\hat{\rho}^{reg}$	0.797*** (0.021)	0.818*** (0.022)	..	..
$\hat{\alpha}^{reg}$	1.911*** (0.188)	1.748*** (0.204)	..	..
$\hat{\rho}^{temp}$	..	..	0.247*** (0.053)	0.262*** (0.048)
$\hat{\alpha}^{temp}$	..	..	0.079*** (0.014)	0.097*** (0.011)
Wald test (F-stat)	91.79 [0.000]	66.33 [0.000]	..	..
R^2	0.587	0.634	0.062	0.070
No. of obs.	1296	1055	1296	1055

Notes: OLS estimation of eq. (3) with $\log Y_{it}^{reg}$ or $\log Y_{it}^{temp}$ as dependent variable. The coefficient $\hat{\rho}^{reg}$ is the estimated persistence for regular income and $\hat{\rho}^{temp}$ is the estimated persistence for temporary income, while $\hat{\alpha}^{reg}$ and $\hat{\alpha}^{temp}$ denote the respective constants. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively. The Wald test shows the F-statistic and in square brackets below the p-value of the null hypothesis that the coefficient of the regular income is equal to one.

Table C.4. Income persistence coefficients for different sub-samples

	(1)	(2)	(3)	(4)
	Regular income	**Regular income**	**Temporary income**	**Temporary income**
	Income below median	Income above median	Income below median	Income above median
$\hat{\rho}^{reg}$	0.846*** (0.030)	0.838*** (0.036)	..	..
$\hat{\alpha}^{reg}$	1.508*** (0.256)	1.535*** (0.341)	..	..
$\hat{\rho}^{temp}$	..	..	0.211*** (0.045)	0.317*** (0.067)
$\hat{\alpha}^{temp}$	..	..	0.084*** (0.011)	0.091*** (0.016)
Wald test (*F*-stat)	26.94 [0.000]	19.73 [0.000]	..	..
R^2	0.447	0.429	0.069	0.068
No. of obs.	1252	1099	1252	1099

Notes: OLS estimation of eq. (3) with $\log Y_{it}^{reg}$ or $\log Y_{it}^{temp}$ as dependent variable. The coefficient $\hat{\rho}^{reg}$ is the estimated persistence for regular income and $\hat{\rho}^{temp}$ is the estimated persistence for temporary income, while $\hat{\alpha}^{reg}$ and $\hat{\alpha}^{temp}$ denote the respective constants. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively. The Wald test shows the *F*-statistic and in square brackets below the *p*-value of the null hypothesis that the coefficient of the regular income is equal to one.

Appendix D. Robustness tests of estimations with lagged shock to temporary income

Table D.1. Sensitivity of total consumption to shocks to income processes of different persistence

	(1)	(2)	(3)
	$\rho^{reg} = 1$ $\rho^{temp} = 0$	$\rho^{reg} = 0.9$ $\rho^{temp} = 0$	$\rho^{reg} = 0.8$ $\rho^{temp} = 0$
ν_{it}^{reg}	0.423*** (0.041)	0.420*** (0.042)	0.392*** (0.042)
ν_{it}^{temp}	0.296*** (0.054)	0.292*** (0.053)	0.282*** (0.051)
ν_{it-1}^{temp}	-0.335*** (0.039)	-0.329*** (0.039)	-0.317*** (0.038)
Wald test (*F*-stat)	5.30 [0.022]	5.36 [0.021]	4.01 [0.045]
R^2	0.119	0.114	0.105
No. of obs.	2351	2351	2351

Notes: OLS estimation of eq. (5) with $\Delta \log C_{it}$ as dependent variable. The variable ν_{it}^{reg} is the regular income shock, ν_{it}^{temp} is the temporary income shock and ν_{it-1}^{temp} is lagged temporary income shock derived from eq. (3). Household size, log of age of household head, and year dummies are included in the estimations but are not shown in the table. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively. The Wald test tests the null hypothesis that the difference between the coefficients of shocks to regular and temporary income is not statistically significant. The values in the square brackets are *p*-values.

Table D.2. Robustness test of the baseline regression to different sets of control variables

	(1)	(2)	(3)	(4)	(5)
ν_{it}^{reg}	0.329***	0.339***	0.334***	0.334***	0.317***
	(0.034)	(0.035)	(0.034)	(0.034)	(0.035)
ν_{it}^{temp}	0.196***	0.200***	0.186***	0.186***	0.183***
	(0.036)	(0.036)	(0.035)	(0.035)	(0.035)
ν_{it-1}^{temp}	-0.247***	-0.249***	-0.237***	-0.237***	-0.234***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
$lage_{it}$	-0.086**	-0.087**	-0.088**	-0.090**	-0.091**
	(0.038)	(0.038)	(0.037)	(0.037)	(0.037)
$\Delta hhsize_{it}$	0.099***	0.097***	0.092***	0.092***	0.094***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
$\Delta partempl_{it}$	..	0.046	0.046	0.046	0.049*
		(0.031)	(0.031)	(0.031)	(0.031)
$\Delta abovec_{it}$	..	..	0.062***	0.062***	0.062***
			(0.016)	(0.016)	(0.016)
$\Delta belowc_{it}$	..	..	-0.090***	-0.090***	-0.089***
			(0.032)	(0.033)	(0.033)
$\Delta realest_{it}$	..	..	..	0.024	0.022
				(0.040)	(0.041)
$\Delta renting_{it}$	..	..	..	0.060	0.064
				(0.057)	(0.057)
$\Delta liquid_{it}$	..	..	..	..	0.015**
					(0.007)
R^2	0.106	0.107	0.119	0.119	0.121
No. of obs.	2351	2351	2351	2351	2351

Notes: OLS estimation of eq. (5) with $\Delta \log C_{it}$ as dependent variable. The variable ν_{it}^{reg} is the regular income shock, ν_{it}^{temp} is the temporary income shock and ν_{it-1}^{temp} is lagged temporary income shock derived from eq. (3). Year dummies are added to the regression but not shown in the table. Robust standard error reported in brackets below the coefficient estimates. Superscripts ***, ** and * indicate that the variable is statistically different from 0 at the 1%, 5% and 10% level respectively.

Appendix 2. Publication II

Reprint from:

Kukk, Merike (forthcoming). How did household indebtedness hamper consumption during the recession? Evidence from micro data. *Journal of Comparative Economics,* available online from 29 July 2015, DOI: http://dx.doi.org/10.1016/j.jce.2015.07.004

Journal of Comparative Economics 000 (2015) 1–23

Contents lists available at ScienceDirect

Journal of Comparative Economics

journal homepage: www.elsevier.com/locate/jce

How did household indebtedness hamper consumption during the recession? Evidence from micro data

Merike Kukk

Tallinn University of Technology, Department of Finance and Economics, Akadeemia tee 3-478, 12618 Tallinn, Estonia

ARTICLE INFO

Article history:
Received 18 July 2014
Revised 7 July 2015
Available online xxx

JEL:
E21
D14
E32

Keywords:
Household indebtedness
Debt repayment burden
Debt-to-income ratio
Amplification effect
Recession

ABSTRACT

Kukk, Merike—How did household indebtedness hamper consumption during the recession? Evidence from micro data

The paper investigates the extent to which household indebtedness suppressed consumption during the economic downturn in 2008–2009. The paper uses a unique quarterly panel dataset containing financial information on over 100 000 individuals. The dataset covers the period 2005–2011, when there were large changes in credit volumes, income and consumption in Estonia, a new EU member country. The estimations show that indebtedness measured by the debt-to-income ratio and the debt service ratio hampers consumption over the whole business cycle. The negative impact of the debt service ratio is, however, substantially stronger during the recession than in the pre-crisis and post-crisis periods, while the negative effect of the debt-to-income ratio is relatively stable over the sample period. The findings suggest that household indebtedness is amplifying the recession and the debt repayment burden indicates the mechanism which is at work. *Journal of Comparative Economics* **000** ( ) (2015) 1–23. Tallinn University of Technology, Department of Finance and Economics, Akadeemia tee 3-478, 12618 Tallinn, Estonia.

1. Introduction

This paper investigates the impact of household debt on consumption during the 2008–2009 recession using administrative data. Household debt has increased extensively, exceeding corporate debt in many countries, and the discussion about the economic implications of household debt has received a great deal of attention in the past decade.[1] A list of studies emphasise the role of household debt in explaining macroeconomic fluctuations and the need to use micro-level data to understand the driving forces behind the business cycles; see the overview by Mian and Sufi (2010a). That paper is one of the few studies to examine the topic with micro panel data.

The issue of the implications of indebtedness is highly relevant in countries in Central and Eastern Europe (CEE), in which household finance started to emerge after the introduction of market-based financial systems in the 1990s. The liberalisation of the financial system led to the entry of foreign-owned banks and this increased access to financial services, including credit products (Beck and Brown 2011). Since then, rapid household credit growth has been observed, and it was particularly rapid in

E-mail address: merike.kukk@ttu.ee

[1] Cecchetti et al. (2011) have highlighted how the total non-financial debt in advanced economies has been rising markedly during the past three decades. Starting at a relatively modest 167% of GDP three decades ago, total non-financial debt has reached 314% of GDP, with the largest contribution to the rise coming from the debt of the household sector, which added 56 percentage points to the increase in the debt.

http://dx.doi.org/10.1016/j.jce.2015.07.004
0147-5967/

new EU member states until the outbreak of the crisis in 2008. To some extent the credit upsurge was justified by the convergence process, but several studies have identified a household credit boom in the CEE countries (Coricelli et al., 2006; Barrell et al., 2009; Herzberg, 2010). Policy institutions were not able to manage the rapid credit expansion in the 2000s, arguably because they lacked previous experience of capital inflows (Hilbers et al., 2005). Consequently, CEE countries suffered severely from the 2008–2009 recession. There is evidently a need to identify the channels through which household indebtedness contributed to the severity of the recession.

The link between household debt and the real economy is household consumption. There are two different strands of theories explaining the effect of household debt on consumption. One strand covers most conventional explanations elaborated from the Life-Cycle/Permanent Income Hypothesis of Modigliani (1954) and Friedman (1957), which note the improved ability to smooth income shocks and so reduce consumption volatility. In the other strand, several policy oriented papers highlight the increased vulnerability of households because of their increased indebtedness and the negative implications of household debt (Barba and Pivetti, 2009; Girouard et al., 2006; Debelle, 2004). A new set of macroeconomic models emphasises the role of credit frictions, where household indebtedness amplifies a recession; see Guerrieri and Iacoviello (2013) and Justiniano et al. (2013) among others.

Although several theoretical studies rely on the amplification effect of household debt to explain recent macroeconomic developments, there are only a few studies that use micro data to investigate the underlying assumptions. Dynan (2012) and Cooper (2012) use the US Panel Study of Income Dynamics (PSID) to analyse the role of household debt in the consumption contraction of households during the last recession. Brown et al. (2012) use the Life in Transition Survey from 2010 to investigate the impact of mortgages on the consumption cutback in European countries. Andersen et al. (2014) examine household leverage and consumption using Danish register data. Given the amount of discussion about the implications of household debt for the economy, there is a great need for microeconomic evidence on the behaviour of indebted households. The current paper addresses this research gap.

The paper contributes to the literature in several ways. First, it uses a unique database which contains quarterly information from 2004:Q4 to 2011:Q4 to investigate the role of household debt in consumption fluctuations. Alongside the financial data, the database includes inflows and outflows from the sight accounts of each individual and this information can be used to proxy income and consumption flows. The quarterly frequency of the data over 7 years is a unique feature of this dataset as no such database has been used before for analysing consumption dynamics. The time period includes periods of rapid economic growth, a very deep recession and the subsequent recovery. On top of that, the number of indebted individuals in the database is over 100 000, which allows different sub-samples to be examined without precision being lost from the estimations.

Second, the paper examines the implications of indebtedness in a small open country, which experienced the largest increase in the household debt burden of any European country in 2000–2007 as the debt-to-disposable income ratio increased from 15% in 2000 to 88% in 2007, while credit tightening was prevalent from 2008 in Estonia.[2] Households in Estonia experienced very volatile income and consumption changes during the last business cycle.[3] Rapid developments in the credit markets and big swings in household income and consumption in Estonia make it possible to investigate the relationship between consumption, income and debt in different parts of the business cycle. Herzberg (2010) tentatively suggests that the decline in consumption in the Baltic countries was partly induced by the debt overhang, meaning that over-leveraged households cut back spending.

Third, the paper investigates the impact of different debt variables on consumption. The relationship between the debt-to-income ratio, the debt service ratio and consumption is less investigated in the literature than the relationship between the debt-to-assets ratio and consumption. All three ratios indicate different mechanisms for how indebtedness spills over into consumption. The results of the paper reveal that the debt-to-income ratio and the debt service ratio have different effects on consumption and that the debt service ratio identifies the channel by which indebtedness amplifies a recession.

The findings can also improve the understanding about the impact of household debt on consumption in other countries that had high rates of debt accumulation in the 2000s. The relationship between indebtedness and consumption growth is expected to be stronger in countries in which the development of financial systems was rapid during the period of economic growth as rapid debt accumulation discloses vulnerabilities. However, as there are only a few studies on the topic available, it will be difficult to draw any distinctions between different European countries. The current estimations serve as a starting point for further investigation of this topic.

The paper proceeds as follows: Section 2 provides a brief overview of the theoretical and empirical literature. Section 3 introduces the hypothesis and the models to be tested. Section 4 introduces the dataset and Section 5 delivers the main features of the variables. Section 6 provides the estimation results. Finally, Section 7 summarises the findings.

2. Literature on the role of household debt in a business cycle

There is a wide range of literature about the relationship between private debt and the business cycle; see Jordà et al. (2011) and the references therein. But there are only a few empirical studies that investigate the role of *household* debt in the business

[2] *Source*: Eurostat database at http://ec.europa.eu/eurostat/data/database. (nasa_10_ki, Gross debt-to-income ratio of households).

[3] The following statistics illustrate the arguments. The average real gross wages increased annually by 13% in 2007 and fell by 5% in 2009. The unemployment rate in Estonia was one of the lowest in Europe at 4.6% in 2007 but by 2010 it was one of the highest at 16.7%. In 2005–2007 real consumption grew annually by 9–13%, but in 2009 it plummeted by 15% from its level of 2008. *Source*: Eurostat database at http://ec.europa.eu/eurostat/data/database. Further details from the author upon request.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

cycle. The seminal paper by King (1994) extends Fisher's debt-deflation theory of an amplification mechanism for the business cycle to the household sector. He relates the deep and long recession in the beginning of the 1990s to the increase in the debt burden during the 1980s. He supports the theory with empirical evidence from ten developed countries for the period 1984–1992 and shows that the slowest recoveries from the global recession in the early 1990s occurred in countries where households had accumulated substantial debt relative to their disposable income.

Similar conclusions have been drawn for the 2007–2009 recession by Glick and Lansing (2010) and Mian and Sufi (2010b), both of which use aggregate data. Glick and Lansing (2010) examine data from 16 advanced countries and find that the countries with the fastest increase in household credit had the deepest declines in real consumption during the recession. Mian and Sufi (2010b) investigate household leverage from 2002 to 2007 across US counties. They find that the counties with the largest increase in the debt-to-income ratio in 2002–2006 showed the largest declines in durable consumption, starting at the end of 2006. The recession began earlier and became more severe in counties with high debt growth than it did in those with low debt growth. These studies identify the negative role of household debt in a recession; however, disaggregated data are needed to provide an understanding of the reasons behind the relationship.

The use of micro level data for studying the implications of household debt for consumption during the last recession is rare. There is one study by Brown et al. (2012) that examines the impact of household debt on consumption in CEE countries, which typically experienced faster debt growth rates than other European countries during the pre-crisis period up to 2008. Brown et al. (2012) use the Life in Transition Survey from 2010. The dataset contains a categorical variable for households' consumption changes. They include in the model a binary variable depicting the presence of a mortgage and find that Eastern European households with mortgages cut back their consumption more than households without mortgages. They do not find such a pattern among Western European households. As the dataset is cross-sectional and includes limited information about liabilities and consumption, it does not fully reveal the behaviour of households.

There are two studies which focus on the relationship between household debt and consumption using the US Panel Survey of Income Dynamics (PSID). Cooper (2012) uses the survey from 1999 to 2009 and compares the different consumption responses to income changes of households with high debt growth and households with low debt growth during the boom. He finds no difference in the consumption response of the two sub-samples. Conversely, Dynan (2012) uses two waves of the PSID survey, from 2007 and 2009, and finds that highly leveraged households with a high debt-to-asset ratio had weaker consumption during the last recession than less leveraged households had. As she finds that high debt also dampened consumption in 2005–2007, it is not clear whether the negative impact on consumption can be related to the recession.

A further study by Dynan and Edelberg (2013) discusses the possible reasons why high indebtedness might hold back consumption beyond the standard wealth effect. The standard wealth effect entails that consumption is positively affected by wealth increases; as debt is negative wealth, the negative relationship between debt and consumption may be induced by the conventional wealth effect. However, they describe several other ways in which leverage may restrain consumer spending, such as by increasing the current borrowing constraints or raising the probability that borrowing constraints will be faced in the future. As indebtedness is taken into account by credit providers when determining credit availability for indebted households, high indebtedness may lead to current or future credit constraints. It is easier to measure the current or actual credit constraints than to identify household's fear of potential credit constraints in the future as the latter is related to household's subjective perceptions. The fear of potential credit constraints is expected to make households cautious in their current consumption decisions.

Another linkage between indebtedness and consumption lies in precautionary motives. Indebtedness entails additional risks to households as they are more vulnerable to negative shocks. Households consider debt servicing to be part of compulsory expenses which they continue to pay under any circumstances (Eurobarometer, 2014). Therefore, high leverage can make households uncomfortable and they may limit consumption for precautionary reasons. Brown et al. (2005) find that unsecured debt has a negative influence on psychological well-being while the qualitative Eurobarometer survey discloses that households experienced negative emotions such as fear about their inability to pay back their loans during the 2008–2009 crisis.

Dynan and Edelberg (2013) use rough proxies for a cut in consumption in the 2009 Survey of Consumer Finance (SCF) to explore the different hypotheses. They find that leverage affects consumption beyond its effect on access to credit. They suggest that high leverage may hinder consumption by increasing households' fears of not having access to credit in the future. However, the results may equally reveal that high leverage makes households uncomfortable, leading them to make conservative consumption decisions.

It has been acknowledged by Dynan (2012) and Dynan and Edelberg (2013) that the survey data have several limitations for investigating the relationship between indebtedness and consumption, such as the large amount of noise, which induces high standard errors that prevent clear conclusions being drawn. Additionally, the relatively long interval between the survey waves makes it difficult to answer questions about particular time periods. This suggests that administrative data should be a valuable source for further research. Baker (2014) uses panel financial data to investigate whether the sensitivity of consumption to income shocks is related to the level of indebtedness. He finds that highly indebted households respond more strongly to income fluctuations, but he does not investigate the relationship between consumption growth and leverage per se.

Andersen et al. (2014) exploit yearly data from several Danish administrative registers from 2003 to 2011 and examine the relationship between household debt and consumption; as the registers do not contain data about consumption, non-housing consumption is imputed from the income and wealth data. They find a strong negative relationship between the pre-crisis

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

loan-to-asset value ratio and a change in consumption during the 2008–2009 recession. As the negative relationship is evident in different socio-demographic, income and wealth groups, they discard the hypothesis that the relationship is induced by current or potential credit constraints. They argue that the increased uncertainty about future financial conditions in 2008 guided highly indebted households to hold back their consumption. However, the tentative conclusion needs further investigation.

This review asserts that there are a limited number of studies on the role of household debt in explaining consumption behaviour. The current study distinguishes between different debt variables to give additional insights into the channels through which household debt spills over into consumption. Households in CEE countries have experienced more severe drops in income and income expectations than households in other European countries, while facing drastic changes in credit conditions at the same time (Brown et al., 2012; Brixiova et al., 2010). This makes Estonia an excellent case for investigating the role of indebtedness in hampering household consumption. Additionally, investigation of the relationship between debt and consumption in different business cycle periods may shed more light on the role of household debt in amplifying a recession. The paper seeks to contribute to the literature by using a time period which covers both the boom and bust phases in the 2000s.

3. Methodology

The paper investigates how much the relative debt level (the debt-to-income ratio) and the relative debt burden (the debt service ratio) affect consumption growth. The debt-to-income ratio is a conventional measure of indebtedness indicating the ability to manage debt over the repayment period as debt is usually paid back from income. The debt service ratio is the ratio of regular interest and principal repayments to income, and measures the ongoing burden of the debt on households as servicing the debt directly affects the current funds available for spending and saving. The two ratios are related as the debt servicing ratio depends on the amount of debt that has been taken. However, two households with the same debt-to-income ratio may have very different debt service ratios if the maturity of the debt is different. An investigation of both variables can help identify the different mechanisms through which indebtedness works on consumption.

A standard consumption function relates changes in consumption to income innovations (Jappelli and Pistaferri, 2010). As households borrow to allocate the use of future income, income expectations increase borrowing. If it is the case that the debt-to-income ratio indicates higher income expectations, a positive relationship between debt-to-income and consumption growth should be observed as consumption decisions take income expectations into account. However, this relationship has not been discussed in the studies which relate household debt to consumption, as a negative relationship has been found between indebtedness and consumption.

As explained in Section 2, indebtedness might relate to future potential credit constraints, meaning that households expect to have difficulties borrowing in order to finance future shocks. This implies that a negative relationship between the debt-to-income ratio and consumption growth should be observed. In Central and Eastern European (CEE) countries the borrowing constraint for households comes mainly from the debt repayment capacity, which is directly related to income, rather than from the value of collateral. This can be explained by the observation that housing equity withdrawal, whereby the home is used as collateral for borrowing for consumption, is more common in countries with advanced financial systems; see Mian and Sufi (2011) for evidence. So a high debt-to-income ratio may indicate a fear of potential credit constraints, as income is an important measure of credit availability.

The same channel may explain the relationship between the debt service ratio and consumption growth, as credit conditions are related to the concurrent ability to service the debt. However, as households are able to choose the duration of the loan which determines the debt service burden, perhaps by extending the repayment period of a loan to lower the debt servicing burden, income is more binding a constraint for future credit than debt repayments are. So if the effect of the fear of potential credit constraints is the main channel through which indebtedness is related to consumption, we would expect the debt-to-income ratio to show a stronger negative effect on consumption than the debt service ratio does.

There is another channel described in Section 2 which can explain the negative relationship between indebtedness and consumption. Highly indebted households should be more concerned about their financial situation than households without debt, because of their larger financial obligations. Larger obligations imply more conservative spending behaviour; this channel can be called the debt distress effect. More servicing of the debt adds an additional burden on households, so the debt distress effect is more explicitly related to the debt service ratio than to the debt-to-income ratio. If the debt distress effect prevails, we should expect it to be the debt service ratio that shows a negative effect on consumption rather than the debt-to-income ratio. An exploration of the two debt variables, the debt-to-income ratio and the debt service ratio, should shed more light on the channels the indebtedness operates through.

The second question is whether household indebtedness has a larger effect on consumption during a recession than at other times in the business cycle. Both of these channels, the concern about potential credit constraints and the distress from the debt burden, are subjective assessments by households and are apparently stronger during a recession. Households are more aware of credit constraints when credit conditions are tighter, so the effect of the fear of potential credit constraints should be different in 2005–2006 and in 2008–2009. Moreover, as households' overall assessment of their current and expected financial situation worsens during a recession, the debt distress is expected to be more prevalent during an economic downturn. This could mean that households might feel more stressed about their debt burden during a recession, leading to a stronger negative relationship between the debt service ratio and consumption growth, ceteris paribus. The strength of the negative relationship between consumption growth and the debt variables during a downturn compared to the relationship in other business cycle periods would indicate the degree of amplification of the recession.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

The paper uses a standard consumption model in which consumption responds to unexpected income changes, see Jappelli and Pistaferri (2010) and Attanasio and Weber (2010).[4] Net wealth can be disentangled into liabilities and assets, and the balance of liabilities has been the main interest in the study of Dynan (2012), Cooper (2012) and Andersen et al. (2014). The liabilities are included in the model in a similar way to Dynan (2012) and Andersen et al. (2014), as lagged variables, assuming that the ex ante debt variables may be associated with consumption growth. The relationship between consumption growth and concurrent debt variables is not of interest because debt is used to finance consumption and the contemporaneous variables should capture the expected positive relationship. The focus of the current paper is on whether the previous debt level or the debt burden can be associated with subsequent consumption changes.

The literature exploring the role of debt in a recession at the aggregate level emphasises the implications of high levels of debt accumulation, suggesting that the change in debt matters rather than that in the level of debt. However, at the micro level the debt level is more informative than the changes in the debt, especially in CEE countries where the opportunities for borrowing emerged relatively recently. If the loan was taken for home purchase which was commonly the case, a sharp increase is observed in indebtedness at the moment of borrowing, which is followed by a slow decline induced by debt repayments at the individual level. In these circumstances a link could be found between the slow de-accumulation of debt and change in consumption but this would not help to explain the negative effect of indebtedness on household consumption behaviour. At the micro level the debt level contains more information about a household's decisions than the debt change does.

In addition to the lagged debt-to-income ratio and the lagged debt service ratio, the lagged asset-to-income ratio is included in the model to capture the effect of liquid financial wealth on consumption growth.[5] The following empirical specification of the consumption model incorporating household wealth components is used:

$$\Delta \log c_{it} = \beta \Delta \log y_{it} + \phi Dtol_{it-1} + \psi Dsr_{it-1} + Z'_{it-1}\alpha + \varepsilon_{it}, \quad (1)$$

where i is an individual person and t denotes time; $\Delta \log c_{it}$ is the change in the log of real consumption and $\Delta \log y_{it}$ is the change in log real income. The coefficient β denotes the sensitivity to income changes, which may also entail information about future income streams. The variable $Dtol_{it-1}$ denotes the lagged debt to income ratio of the previous period and the variable Dsr_{it-1} stands for the lagged debt service ratio of the previous period. The coefficients ϕ and ψ depict the effect of the respective ratios on consumption changes. The vector of Z_{it-1} contains the other balance sheet components, such as the lagged asset-to-income ratio, and other control variables that may also affect consumption changes.

There are two issues related to the current consumption model. First, Eq. (1) is not the complete consumption model as the model obviously does not include all the determinants of consumption change such as preferences and real wealth. These might potentially be correlated with income and debt related variables, hence the error term ε_{it} is expected to be correlated with the explanatory variables.

Second, there is a substantial number of households who do not own any debt, but it is only possible to estimate the coefficients ϕ and ψ in Eq. (1) for indebted households. Debt ownership is affected among other things by the attitude towards debt or the willingness to borrow, as discussed by Chien and Devaney (2001). It is reasonable to assume that unobserved characteristics which affect debt ownership may also affect the consumption change of households. For example, if indebted households are more eager to change their consumption because of preferences which are not observed, the debt variable would pick up this effect and the estimated ϕ and/or ψ will be biased. In this case selection bias needs to be taken into account.

With a panel dataset, a fixed effects model directs both the endogeneity problem and the selection bias and therefore a fixed effects model is used:

$$\Delta \log c_{it} = u_i + \beta \Delta \log y_{it} + \phi Dtol_{it-1} + \psi Dsr_{it-1} + Z'_{it-1}\alpha + \tau_t + \varepsilon_{it}. \quad (2)$$

In Eq. (2) the term u_i is the individual fixed effects and the term τ_t the time fixed effects. The individual fixed effects partial out the effect of time-invariant omitted variables such as preferences. The time fixed effects are included to control for time-varying heterogeneity that may stem from omitted common variables or global shocks such as wealth changes induced by aggregate real estate price movements. The current model specification does not control for the endogeneity induced by the effect of the dependent variable on the explanatory variables, such as the wish to increase consumption, which may lead to additional borrowing and an increase in indebtedness. However, as *lagged* explanatory variables are used rather than contemporaneous ones, this particular source of endogeneity is removed.

The main focus of the paper is on household debt, and the coefficients ϕ and ψ show the extent to which the lagged debt variables are associated with consumption growth. If indebtedness dampens consumption growth, the estimated coefficients ϕ and/or ψ should be negative, as households which are indebted at a higher level constrain their consumption growth more than the same households with a low debt level would, ceteris paribus. The current model does not relate the effect of indebtedness to

[4] The standard consumption model assumes that all the predicted changes in income are already taken into consideration when consumption decisions are made. Only unexpected changes in current income or new information about future income should affect consumption. However, current income changes contain new information about future income levels if income changes are relatively persistent, as shown in the study of the persistence of income shocks by Kukk et al. (2015). As the focus of the paper is on the relationship between consumption and balance sheet variables, income change is not decomposed into expected and unexpected income shocks. This simplification should not alter the results of the variables being observed.

[5] The ratios are used rather than gross volumes in order to keep the observations where the value of the variables is zero in the estimations. If the variables were in logs, which would be an alternative approach, these observations would be lost.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

any other variable which determines the consumption change but estimates the impact of indebtedness on consumption growth per se.[6]

The debt-to-income ratio may be high because the debt volume is high or because income is low. It is not relevant to distinguish the role of the two components as the *relative* debt level is important when individuals assess their capacity to repay their debt, as the debt is usually repaid from income. The debt ratio can increase either because the debt is increasing in size or because income is decreasing. The same applies to the debt service ratio.

The current model specification does not rule out the possibility that the high debt level or high debt service ratio will affect not only the consumption change but also the other choices of a household such as the labour supply, which would affect the income change. In this case the income change and the debt variables in Eq. (2) are correlated and the estimated coefficients ϕ and ψ show only the direct effect of the indebtedness on the consumption change, not taking into account the indirect effect via the income change. The relationship between indebtedness and the labour supply is beyond the scope of this paper.

Furthermore, the income change in Eq. (2) may be endogenous as some of the income changes are anticipated and therefore already incorporated into the consumption decisions. This would lead to a lower estimated coefficient for the income change. This would be a concern in analysing the interactions between income and consumption, but when examining the effect of indebtedness on consumption per se it is important to control for both unanticipated and anticipated income changes. Indebtedness is expected to be positively correlated with future income change as the life-cycle model suggests that households borrow in anticipation of income increases. This makes it necessary to control for the change in total income, which captures both the anticipated and unanticipated components of income change. If the anticipated income changes are not controlled for in Eq. (2), the estimated coefficients to the debt-related variables would pick up a positive relationship between anticipated income change and consumption resulting in an upward biased coefficient for debt related variables.[7]

The parameters ϕ and ψ in Eq. (2) can be estimated for different time periods in order to examine the stability of the estimated coefficients. If indebtedness amplifies the recession, the relationship between the lagged debt-to-income ratio or the lagged debt service ratio and consumption change should be different in the period of economic growth and in the recession. More precisely, the amplification effect implies that the negative coefficient ϕ or ψ is lower during the recession (larger in numerical terms) than when the economy is growing. Previous papers in the area do not investigate the dynamics of the impact of indebtedness on consumption over the business cycle. Therefore the following model is also estimated:

$$\Delta \log c_{it} = u_i + \beta_t \Delta \log y_{it} + \phi_t Dtol_{it-1} + \psi_t Dsr_{it-1} + Z_{it-1}\alpha_t + \tau_t + \varepsilon_{it}, \tag{3}$$

where the estimated coefficients β_t, ϕ_t, ψ_t and the vector of the estimated coefficients α_t may vary across time periods. In addition to the assumption of time-varying coefficients of ϕ or ψ, it is reasonable to relax the assumption of the constant parameter of β and the vector of constant parameters of α. The sensitivity of consumption to an income change depends on the persistence of the income change, which might vary over the business cycle. Additionally, the effect of assets and repayment problems is allowed to vary. The paper does not focus on the changes in these coefficients, but a relaxation of the assumption about constant coefficients is needed to avoid the estimated coefficients ϕ and ψ picking up the effect of other time-varying coefficients if these are constrained to be constant.

4. The dataset

The paper uses a unique anonymised client database from a financial institution. The database contains quarterly information about individuals who were regular clients of the financial institution from 2004:Q4 to 2011:Q4.[8] The panel structure of the database makes it possible to follow the changes in the behaviour of the same individual over a sequence of time periods. It is therefore possible to compare the effect of debt on consumption over the business cycle.

Although the database contains information from only one financial institution and does not cover all the financial information about an individual, this is not considered to be a major concern. In Estonia there is very high financial inclusion of individuals, as over 90% of individuals own an account at a formal financial institution while over 80% own a card, and these shares are among the highest in Europe (Demirgüç-Kunt and Klapper, 2012; Beck and Brown, 2011). On top of that, the financial sector is very concentrated in Estonia and most individuals make all their financial transactions in one financial institution.

The main limitation of the dataset is that it contains information at the individual level while consumption behaviour is conventionally analysed at the household level. A household may have a joint bank account but in most cases the income and expenditure of a household are split between the bank accounts of household members. As the database does not allow household members to be matched, some additional robustness tests are warranted.

An important feature of the dataset is the information about quarterly inflows from legal entities (companies or other institutions) and outflows from the sight accounts of an individual. As mentioned earlier, a majority of the financial transactions of individuals are made through sight accounts, and payments from legal entities can be interpreted as earnings. Alternatively, the total inflow to a sight account could be considered as earnings but this variable is considered to be endogenous as each individual

[6] Although the consumption sensitivity to income changes is an important research question and the topic has been investigated by Baker (2014), this analysis has been disregarded to keep the focus of the paper.

[7] The standard consumption theory suggests that households do not respond to anticipated income changes. Nevertheless, empirical evidence show that anticipated income changes affect consumption, see the overview by Jappelli and Pistaferri (2010).

[8] Regular bank clients are the clients who have regular inflows to and outflows from their sight accounts.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Table 1
Description of the main variables in the dataset.

Variable	Definition
$\log y_{it}$	Logarithm of real yearly inflow from legal entities to sight accounts of individual i in quarter t, in EUR in 2005 prices
$\log c_{it}$	Logarithm of real yearly outflows from sight accounts of individual i in quarter t, excluding transactions between saving and investment accounts, in EUR in 2005 prices
$DtoI_{it}$	Debt-to-yearly income ratio; debt stock is measured at the end of quarter t and income is the sum of the income of the four previous quarters
Dsr_{it}	Ratio of annual debt service payments to annual income in quarter t

can determine the payments to the sight account which are not made by legal entities. These payments are mainly payments from savings or sight accounts at other financial institutions or transactions between spouses and relatives. Therefore the inflow from legal entities is used, which contains such income sources as salaries, benefits, dividend payments and self-employment activities. The proxy excludes other official or unofficial income. Similarly, as shown by Kukk and Staehr (2014) the survey data contains gaps in the reported income.

The outflows from the sight account exclude payments to savings or other accounts held by an individual that reflect saving decisions. The outflows also exclude payments to the commercial bank that mainly consist of interest and principal payments on the debts. When these transactions are excluded, the rest of the outflows from the sight account can be interpreted as spending. The outflow from the sight account can take the form of payments, transactions or cash withdrawals. There is still some measurement error for the spending variable as the outflows from the sight account do not include imputed rent and the spending on housing is therefore lower for a homeowner than for a renter. As the paper focuses on *changes* in inflows and outflows, the differences in levels do not affect the changes in spending as long as the tenure status does not change during the period, or if it is controlled for by tracking of homeownership. Additionally, the outflows from the sight accounts contain real estate purchases as it is not possible to distinguish the purpose of the payments from the sight accounts. However, it is possible to control for real estate purchases if these are made using a loan. The dynamics of mean inflows and outflows from the sight account are compared to the aggregate statistics in the next section.

In order to control for seasonal fluctuations and make the variables comparable to standard measures for households, yearly income values are computed. Yearly income is the sum of inflows from legal institutions to a sight account in the current quarter and the three preceding quarters. The outflow from the sight account (excluding payments to saving accounts) in the current quarter and the three preceding quarters is used as a proxy for yearly consumption. The yearly consumption and income variables are computed for each quarter.

The database contains information on the stock of total financial liabilities of the individual, which comprises the balance of housing loans, consumer loans, overdrafts and revolving credit cards at the end of each quarter. The database includes a flag for any new or any additional housing loan in each quarter. As noted, this is needed to control for outflows from the sight account for purposes other than consumption and for changes in homeownership.

The two variables of indebtedness that are used in the paper are the debt-to-income ratio and the debt service ratio. The first ratio is computed by dividing the stock of debt at the end of the period by the income from legal entities for the preceding year (four quarters). Yearly income is used instead of quarterly income to offset seasonal changes of income, which would lead to seasonal volatility of the ratio as the debt stock is relatively stable across a year.

The second variable related to debt, the debt service ratio, is computed as debt payments in the current and three previous quarters divided by the income of the same quarters. The yearly ratio is computed to offset the short-term fluctuations of income. The nominal variables of consumption and income are deflated by the HICP consumer price index and the real variables are expressed in 2005 prices; the variables are expressed in logarithms.

The notation of the main variables in the model is given in Table 1. The panel summary statistics of the main variables are given in Table A.1 of Appendix A.

The dataset contains a variable depicting the stock of financial assets including deposits, funds, bonds and shares at the end of each quarter. It is possible to control for the differences in consumption changes between individuals with different financial assets. The dataset lacks information on the real estate of the individual, but it has been argued by Buiter (2010) that changes in home values do not change the wealth of households and it is posited by Attanasio et al. (2009) that the wealth effect on consumption may actually capture the effect of common underlying factors which are controlled for by the time fixed effects in the model.

The dataset contains information about the ownership of different saving products such as pension insurance and life insurance that imply regular outflows of savings from sight accounts. These variables are used to control for the outflow from the sight account to the saving products. The dataset includes some demographic variables such as age and gender, and these are used for robustness checks across different sub-samples. The main description and statistics of the additional control variables is given in Table B.1 of Appendix B.

There are in total 2 597 000 observations in the pooled dataset across 29 quarters from 2004:Q4 to 2011:Q4 for 108 000 individuals or around 12% of the working-age population of Estonia. The initial filtering of the total client database has been done by the financial institution, excluding all individuals below the age of 20 in 2011:Q4 and above the age of 70 in 2004:Q4. In order

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

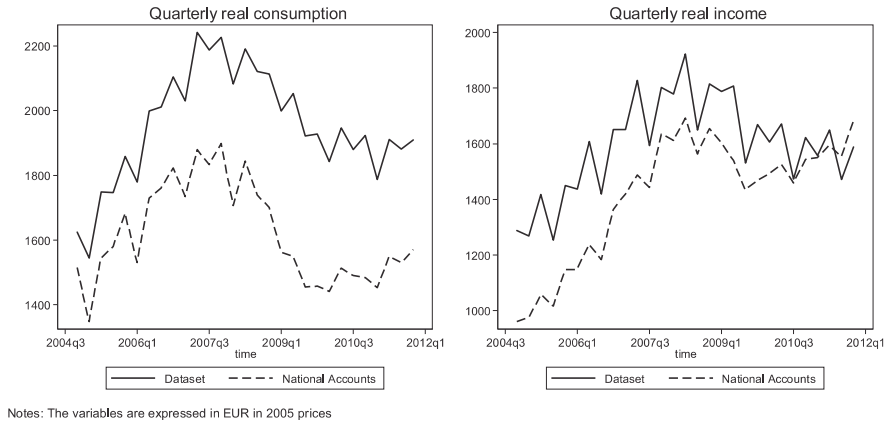


Fig. 1. The dynamics of quarterly consumption and income from 2004:Q4 to 2011:Q4.

to avoid double banking relationships, the financial institution has excluded all private clients for whom the financial institution is not identified as the main bank for the whole time period of 2004–2011.

Some additional exclusions from the dataset have been applied, such as the exclusion of all individuals who could be identified as farmers, self-employed, entrepreneurs or family doctors as there is evidence that these groups use their private accounts for business purposes. Additionally, the individuals for whom any of the quarterly inflows or outflows of the sight account fall in the upper 99th or 100th percentiles have been excluded. Individuals whose transactions of securities (funds, bonds and shares) fall into the 1st, 2nd, 99th or 100th percentiles were also excluded; extraordinarily large transactions from sight accounts have a higher probability of denoting a transaction atypical of the income or expenses of an individual. Observations with other abnormalities such as negative values for any original variable, or with a debt-to-income ratio or debt service ratio in the 99th or 100th percentile were also excluded.

5. Description of the variables and data validation

Although the dataset covers about 12% of the working age population, which is far more individuals than any household survey, additional robustness analysis has been done to validate the representativeness of the data. The summary statistics for the main variables are given in Table A.2 of Appendix A. As the consumption in the dataset is proxied by outflows from the sight account, the variable is compared to the household aggregate per capita consumption.[9] The dynamics for the mean real spending of an individual in the dataset and per capita consumption of the household sector from the National Accounts from 2004:Q4 to 2011:Q4 are presented on the left-hand graph in Fig. 1.

The dynamics of the two variables are very similar, and the only difference is that the mean consumption in the dataset is somewhat larger than per capita consumption from the National Accounts. There may be several reasons for this difference. First, the dataset does not contain individuals younger than 20 or older than 70, and these population groups exhibit lower average consumption than the population group between 20 and 70. Second, the average income in the sample might be higher than the average income in the population as the population group which does not use the services of financial institutions is usually the low-income group. And third, the proxy for consumption in the dataset contains real estate purchases as it is not possible to identify the outflows from the sight account for this particular purpose. However, as mentioned in Section 3, it is possible to control for the real estate purchases with a housing loan in the model.

The comparison of the quarterly dynamics of the income variable in the dataset and in the National Accounts is given on the right-hand graph in Fig. 1. The income is proxied by the mean of inflows to the sight account in the dataset and per capita disposable income of the household sector is used from the National Accounts. The average income according to the dataset is somewhat higher than the aggregate per capita disposable income. One reason for this is that the sample is drawn from the working age population, which earns a higher income than other population groups like the retired population. Additionally, the mean income in the dataset shows stronger seasonality than the average aggregate income. Again, the reason might be related to stronger seasonality in the earnings of the working age population compared to the earnings of the retired population.

[9] Total final consumption of the household sector from the National Accounts is divided by the total population. All the aggregate statistics in Section 5 are taken from Statistics Estonia http://www.stat.ee/database and the Bank of Estonia http://www.eestipank.ee/en/statistics. Further details are available from the author upon request.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

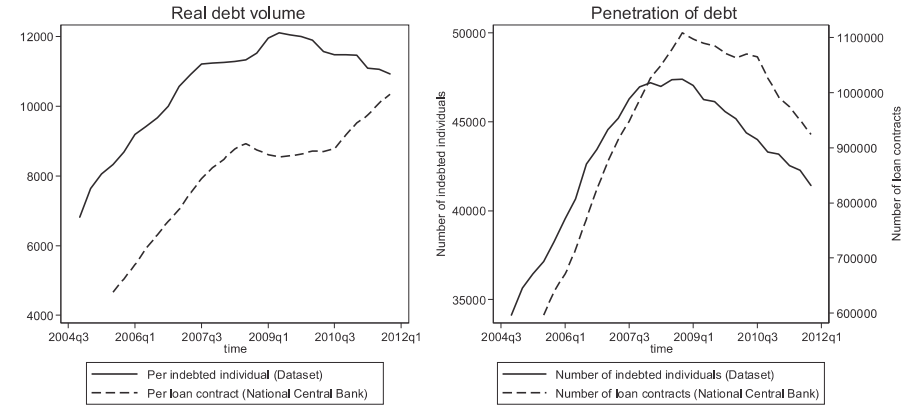


Fig. 2. The dynamics of mean and median loan volumes for total debt and the number of individuals with debt from 2004:Q4 to 2011:Q4.

Nevertheless, the dynamics of the two income variables are very similar and the evolution of the variables differs only in 2011. The upshot of the analysis is that the dynamics of average consumption and income variables are very similar in the dataset and in the aggregate statistics, suggesting that the dataset reflects the economic situation of households relatively well.

Next, the dynamics of real debt volumes and the penetration of debt are compared to the aggregate statistics from Eesti Pank, the central bank of Estonia. Currently no survey data on financial assets and liabilities are available in Estonia and therefore aggregate statistics on household debt volumes and loan contracts have to be used for comparison. The left-hand graph in Fig. 2 gives the evolution of the mean value of the total real debt balance of an indebted individual in the dataset from 2004:Q4 to 2011:Q4. It is compared to the average real value of the loan contract. As an individual may hold more than one loan contract, the average debt balance of an individual is higher than the average value of a loan contract. The average debt balance increased rapidly during 2007–2008 and peaked in 2009:Q2 but then started to decrease. Housing loans contribute most to the total loan volumes in Estonia, and 80% of the total loan volume was in housing loans at the end of 2004 according to the statistics of Eesti Pank, a share that had increased to 83% by the end of 2011.

The right-hand graph in Fig. 2 presents the number of individuals who own any type of debt, whether as housing loans, consumer debt, or the balance on credit cards or overdrafts. The number of indebted households increased very fast during 2004–2007, after which there was a decline between 2008 and 2011. The number of loan contracts shown in data from Eesti Pank shows very similar dynamics. The decline in indebted households can be explained by a substantial drop in households' demand for credit during the recession, as found by Meriküll (2015). At the end of 2011, about 40% of households in the sample were indebted, indicating that the behaviour of indebted households would have a major impact on the aggregate behaviour of the household sector.

Additionally, the quarterly dynamics of the debt-to-income ratio are presented in Fig. 3 and the quarterly statistics are given in Table A.2 of Appendix A. On the left-hand graph in Fig. 3 the debt-to-income from the National Accounts and the dataset are compared. As the debt-to-income ratio is computed for the total population or total sample, it does not take into account the changes in the penetration of debt as seen on the right-hand graph in Fig. 2. According to the National Accounts, debt-to-disposable income increased more during the period than the average debt-to-income ratio in the dataset. The differences can be explained by the higher average income in the sample than in the total population. However, the differences suggest that additional robustness tests need to be done for different income groups to understand the implications of the differences between the aggregate and sample statistics.

The right-hand graph in Fig. 3 presents the debt-to-income ratio in the dataset when conditioned on debt ownership. This variable reflects the actual indebtedness of individuals taking into account the penetration of debt. This statistic is not available from the aggregate statistics, so the sample statistics are analysed. The mean debt-to-income ratio was relatively stable during the period of economic growth, increasing slightly from 1.18 in 2005:Q3 to 1.28 in 2007:Q3 due to the rise in the debt volumes, and falling slightly to 1.23 by 2008:Q3. A substantial increase in the debt-to-income ratio occurred from 2008:Q3 to 2010:Q2 when the ratio reached 1.42. The increase was mainly induced by the decline in household income, while the mean debt volume was relatively stable (see Figs. 1 and 2). Since the beginning of 2011 the debt-to-income ratio has decreased slightly.

Fig. 4 presents the dynamics of the debt service ratio while the quarterly statistics are provided in Table A.2 of Appendix A. There are no public statistics about the debt service ratio, neither from any survey nor from aggregate statistics. It is apparent that the debt service ratio of *individuals* found here is higher than the actual debt service ratio of *households* as the debt is usually

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

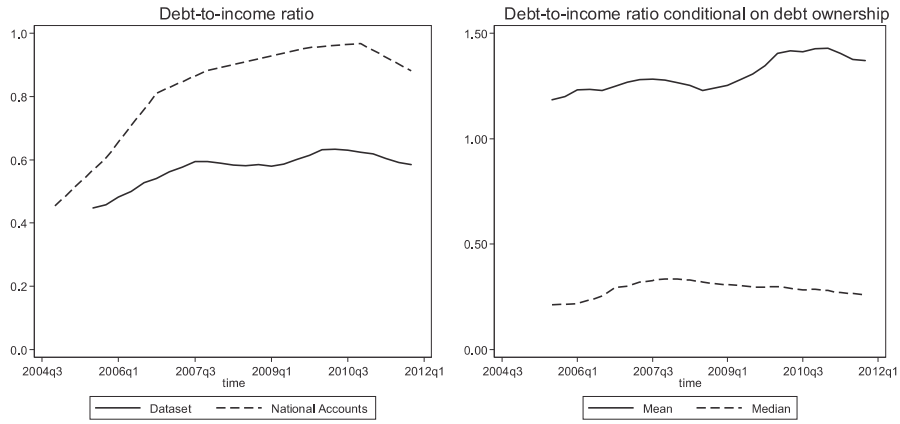


Fig. 3. The dynamics of the debt-to-income ratio 2005:Q3–2011:Q4.

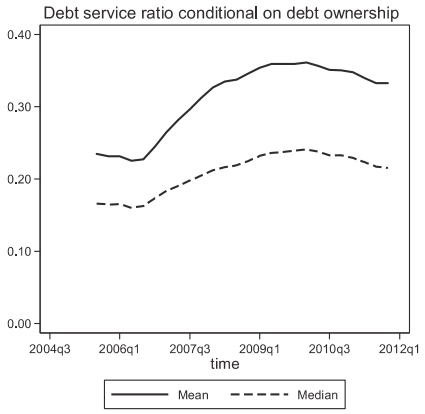


Fig. 4. The dynamics of the mean and median debt service ratios 2005:Q3–2011:Q4.

serviced from one current account which belongs to one household member. However, when the *relative* burden across segments and the *changes* in the burden over the business cycle are compared, the differences between individuals and households are expected to be less pronounced. The evolution of the debt burden is expected to be relatively accurate given that the income and debt variables follow the dynamics of the aggregate statistics.

The debt service ratio shows larger changes than the debt-to-income ratio does during the sample period. The mean debt service ratio increased from 0.23 in the middle of 2006 to 0.33 in the middle of 2008. The following period showed a modest increase in the ratio to 0.36 by 2010:Q1 and a slight decrease since then. Hence, the debt payment burden of individuals increased even more rapidly when the economy was growing than it did during the recession, and the reason behind the change in the ratio varies over the period. The increase during the period of economic growth was induced by the increases in average debt volumes as given in Fig. 2. The increase during the recession was caused by the decline in income as shown in Fig. 1.

6. The estimations

6.1. The empirical model

Due to the special features of the database, some adjustment is needed to Eq. (2). As shown in Fig. 1, there are substantial seasonal fluctuations in the quarterly data for consumption and income, and therefore the income and consumption variables for

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

each quarter are computed on a rolling basis where four previous quarters are added together, as explained in Section 4. In order to estimate the yearly change in consumption for two consecutive years, the difference in log consumption between quarters t and $t-4$, $\Delta_4 \log c_{it} = \log c_{it} - \log c_{it-4}$, has been used. The yearly change in income is expressed as the difference in log income between t and $t-4$: $\Delta_4 \log y_{it} = \log y_{it} - \log y_{it-4}$. The debt variables are lagged by four quarters, $Dtol_{it-4}$ and Dsr_{it-4}.

The following equation, an extended version of Eq. (2), was estimated:

$$\Delta_4 \log c_{it} = u_i + \beta \Delta_4 \log y_{it} + \phi Dtol_{it-4} + \psi Dsr_{it-4} + Z'_{it-4}\alpha + X'_{it}\gamma + \tau_t + \varepsilon_{it}. \quad (4)$$

The vector Z' contains wealth-related and debt-related variables which might affect consumption growth, such as the lagged ratio of financial assets to income and a lagged dummy for debt repayment problems. The vector X' contains other control variables, which are needed because consumption is a proxy for the outflows from the sight account. A brief description of the control variables required follows.

First, outflows from sight accounts include transfers to saving products offered by other financial institutions. These outflows should not be considered as expenditure but as saving. Transfers to investment and savings accounts in the same financial institution have ex ante been deducted from the outflow of sight accounts. It is also possible to control for regular saving in the insurance products of the financial institution by including a dummy for acquiring a life insurance or a pension insurance product.

Second, the servicing of a housing loan used for home purchase should not be counted as consumption, even though these funds are an outflow from the sight account. So when the outflow from the sight account for a period when a housing loan was taken is compared to the outflow for the following period, a significant decrease in the outflow from the sight account should be observable. On top of that, there might be an additional effect from the housing loan when the individual changes tenure status from renter to homeowner as rent payments are replaced by debt payments. Rent payments are reported as outflows from the sight account while debt payments are not. This might lead to estimated negative coefficients ϕ and ψ even if housing costs remain the same and rent payments equal debt payments while the consumption of other goods is kept the same. Therefore any new or additional housing loan that would alter the results should be taken into account. Initial estimations showed that a new or additional housing loan increases the outflow from the current account in the same and the following quarters. For this reason, dummies for new or additional debt are added for the time periods between t and $t-8$. The full list of control variables is given in Table B.1 of Appendix B.

The fixed effects model assumes stationarity of the variables as the use of a fixed effects model with a non-stationary series of variables would lead to spurious regression. Most unit root tests assume an asymptotic time property and without a sufficiently long time period the unit root tests are weak. The unit root tests of Harris and Tzavalis (1999) and Im et al. (2003) can be used for samples with a fixed time period and an asymptotic sample size, which are the features of the dataset used in this study. The Harris–Tzavalis test assumes that all panels share the same autoregressive parameter while the Im–Pesaran–Shin test relaxes the assumption of a common autoregressive parameter. The Harris–Tzavalis test requires a strongly balanced dataset, while the Im–Pesaran–Shin test does not require a balanced dataset, though there cannot be gaps within a panel. Therefore both tests can be implemented only for data series which do not contain missing observations. Both unit root tests take cross-sectional dependence into account.

Empirical studies using aggregate data typically find a unit root and co-integration for household log real consumption and log real income, but $I(0)$ for the variables in differences; see Morley (2007) and the references therein or Aben et al. (2012) for the Estonian aggregate data series. Additionally, empirical evidence suggests that the data series of aggregate household debt and assets may be non-stationary; see Bassanetti and Zollino (2010) and Martinez-Carrascal and Rio (2004). Therefore the unit root tests were implemented for income, consumption and all wealth-related ratios. The results of the Harris–Tzavalis and Im–Pesaran–Shin tests are given in Table B.2 of Appendix B. Both unit root tests reject the hypothesis that log income change and log consumption change, or the debt-to-income ratio, the debt service ratio and the asset ratio, contain a unit root. The unit root test results are consistent with the results of other studies and the tests suggest that the current model specification in Eq. (3) can be estimated by fixed effects.

6.2. Estimations for the full sample period

The main interest of the paper is the coefficients ϕ and ψ, which express the impact of the *previous* debt-to-income ratio and debt service ratio respectively on *current* consumption decisions. The coefficients are expected to be negative under the hypothesis of the negative impact of debt on consumption as discussed in Section 3. The estimations cover the period from 2006:Q4 to 2011:Q4.[10]

The results of the baseline estimations of Eq. (4) for the full period using two different debt variables are given in Table 2. The estimated coefficients of the main variables are given in the table while the results for the full model with the control variables are given in Table B.3 of Appendix B.

The estimations reveal that the hypothesis about the negative impact of debt consumption is confirmed. Table 2 column (1) shows that the coefficient for the lagged debt-to-income ratio ($Dtol_{it-4}$) is −0.048, implying that the average debt-to-income ratio of 1.3 in the sample is associated with a yearly consumption decline of 6.1%. Fig. 3 shows that between 2008:Q3 and 2010:Q2

[10] Although the database starts from 2004:Q4, the yearly income and yearly consumption variables set the earliest starting time at 2005:Q3. The inclusion of $t-8$ lagged control variables means that the first time period in the estimations is 2006:Q4.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Table 2
Coefficient estimates for the full sample period from 2006:Q4 to 2011:Q4.

	(1)	(2)	(3)
$\Delta_4 \log y_t$	0.5460***	0.5447***	0.5491***
	(0.0035)	(0.0035)	(0.0035)
$Dtol_{it-4}$	−0.0478***		−0.0367***
	(0.0016)		(0.0016)
Dsr_{it-4}		−0.1792***	−0.1397***
		(0.0038)	(0.0040)
Other control variables	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
R^2	0.2803	0.2804	0.2819
No. of groups	102 968	102 968	102 968
No. of obs.	1 733 332	1 733 332	1 733 332

Notes: FE estimation of Eq. (4). Control variables and time dummies are included in the estimations but not shown in the table. Standard errors are reported in parentheses below the coefficient estimates, SE estimates are robust to disturbances that are heteroskedastic and autocorrelated. ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

the debt-to-income ratio increased by 0.2 points in the sample, suggesting that during this period the negative effect on yearly consumption change was slightly stronger than it was before 2008.

When Eq. (4) is estimated with the lagged debt service ratio, Dsr_{it-4}, the results are quite similar. Table 2 column (2) shows the coefficient for the lagged debt service ratio to be −0.18, implying that the average debt service ratio of 0.31 in the sample is associated with a yearly consumption drop of 5.6%. As the debt service ratio increased between 2006:Q3 and 2010:Q1 by 0.13 points, the suppression of consumption was presumably stronger in 2009 than in 2006.

The third set of estimations includes both the lagged debt-to-income ratio and the lagged debt service ratio in one model to see whether they capture the same effect of indebtedness or whether there are different mechanisms at work as pointed out in Section 3. Table 2 column (3) gives the results for the model with two debt variables. The estimated coefficient of the lagged debt-to-income ratio is −0.037 and the coefficient of the lagged debt service ratio is −0.14. Both variables remain statistically significant, though slightly smaller than in columns (1) and (2), which indicates that the debt variables capture the same mechanism to some extent but both variables carry additional information. When taking into account the average values of the variables, which are given in Table A.1 of Appendix A, the estimations suggest that the debt-to-income ratio can be associated with a drop in yearly consumption of 4.8% and the debt service ratio can be associated with a decline in yearly consumption of 4.3%.

Table B.3 columns (1)–(3) in Appendix B provide the estimation results of Table 2 for all variables including the control variables. Columns (4) and (5) in Table B.3 of Appendix B show the estimations without time fixed effects and without individual effects respectively. The exclusion of time fixed effects alters the estimated coefficient for the debt service ratio, indicating that this variable would pick up negative aggregate shocks when time dummies are not included. When individual heterogeneity is not taken into account, both debt variables are upward biased, meaning the coefficients are less negative, which is consistent with the theoretical assumptions as modelled by Eggertson and Krugman (2012). These are that if borrowing households are less patient than households who do not borrow, the cross-sectional variation will show a smaller negative relationship between consumption growth and indebtedness.

If indebtedness leads to debt repayment problems, individuals will dampen their consumption to improve their financial situation. The current estimations do not pick up the effect of debt repayment problems as the dummy for debt repayment problems is included in Eq. (4). The estimated coefficient is provided in Table B.3 of Appendix B; the dummy is negative and significant. Additionally, the effect of liquidity constraints is controlled for by the inclusion of the lagged financial asset to income ratio in the model. As expected, the estimated coefficient of the financial asset to income ratio is positive (see Table B.3 in Appendix B).

Table B.4 in Appendix B provides additional robustness tests by excluding the other control variables one by one, which are income change (column (1)), financial assets (column (2)), a dummy for debt repayment problems (column (3)), dummies for saving products (column (4)), and dummies for a new or additional housing loan (column (5)). When the income variable is excluded, the explanatory power of the model decreases substantially, from 28% to 11%. The outcome of the exclusion of the main variable in the consumption model is that the estimated coefficients for the debt-to-income ratio and the debt service ratio are upward biased as the variables are positively correlated with income change and they pick up the positive relationship between income and consumption change, as explained in Section 3. Additionally, the estimations of the debt-to-income and debt service variables are somewhat affected by the exclusion of the housing loan dummies, while the estimated coefficients are robust to the exclusion of the other variables. The overall result of the robustness estimations is that the results are robust to different model specifications.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

The upshot is that indebtedness is negatively associated with changes in household consumption; households lower their consumption when their debt levels increase. The estimations confirm the hypothesis that debt has a negative effect on consumption and are consistent with other studies; see Andersen et al. (2014), Dynan (2012), and Brown et al. (2012). Only Dynan (2012) includes the debt service ratio in the model, showing that both leverage (the debt-to-asset ratio) and the debt service ratio are negatively related to consumption growth in the period 2007–2009. The estimations confirm that both debt-to-income and the debt service ratio are important for the consumption decisions of households, so both the relative indebtedness and the current debt burden matter. If there are two individuals with the same debt-to-income ratio, the one who has higher debt repayments because the maturity of the loan is shorter or the interest rate is higher will suppress consumption more than the individual with low debt payments.

6.3. Quarterly estimations

The estimations for the full period do not give any clear picture about the possible mechanism through which indebtedness constrains consumption, they only suggest that there are different mechanisms at work as different debt variables are significant in the model. However, as the effects of the debt-to-income ratio and the debt service ratio on consumption change exhibit similar magnitude, it is not obvious which mechanism is at work. Having different estimations for the period of economic growth and for the recession would shed more light on the mechanisms.

If consumption is similarly dampened during the period of economic growth and the recession, then the indebtedness would not amplify the recession. The amplification effect of debt implies that the negative effect on consumption is stronger during an economic downturn than during the other parts of the business cycle. The effect of the fear of potential credit constraints should be stronger during a recession when credit conditions tighten. Similarly, the debt distress effect might be stronger during a recession following the overall sentiment of households.

It is usually difficult to assess the change in the effect as frequent data for a long time period are needed. Andersen et al. (2014) use yearly panel data and estimate the model for different time slots, with the initial time fixed at 2007, meaning the relationships between the leverage in 2007 and the consumption changes in 2007–2008, 2007–2009, 2007–2010 and 2007–2011 are estimated. The estimations provide a cumulative effect and the estimated negative parameter for leverage doubles from 2008 to 2010 but remains the same in 2011. Rolling estimations for each time period would provide additional information about the change in the effect of debt on consumption growth.

The dataset used here contains so many quarterly observations that it is possible to estimate the coefficients ϕ and ψ for each quarter separately. The model including both the debt-to-income ratio and the debt service ratio is preferred, otherwise one variable may capture the effect of the omitted variable as shown in Table 2, and this would make the interpretation of the results more difficult. The model does not suffer from multicollinearity problems as the correlation between the two debt variables is modest at 0.48 and the number of observations is very large.

The following extended model of Eq. (3) was estimated:

$$\Delta_4 \log c_{it} = u_i + \beta_t \Delta_4 \log y_{it} + \phi_t Dtol_{it-4} + \psi_t Dsr_{it-4} + Z_{it-4}\alpha_t + X'_{it}\gamma + \tau_t + \varepsilon_{it}. \quad (5)$$

In this model the coefficient ϕ_t denotes the impact of the lagged debt-to-income ratio on consumption change and the coefficient ψ_t denotes the impact of the lagged debt service ratio on consumption change in quarter t. It is not only the coefficients for debt variables that are allowed to vary quarterly but all other coefficients are too. All other control variables and time fixed effects in Eq. (3) are included in the model.

The estimated quarterly coefficients for the lagged debt-to-income ratio and for the lagged debt service ratio are given in Table C.1 of Appendix C. The dynamics of the estimated quarterly coefficients are captured in Fig. 5, where the quarterly point estimates are presented together with the 95% confidence intervals.

Although the estimations for the full period given in Table 2 indicate quite a similar magnitude for the effects of the lagged debt-to-income ratio and the lagged debt service ratio on consumption growth, the quarterly estimations reveal differences. The negative relationship between the lagged debt-to-income ratio and consumption change fluctuates around −0.05 in 2006–2008 and decreases afterwards. The point estimate is −0.053 for 2006:Q4 and −0.052 for 2008:Q4, indicating that individuals with a 0.1 point higher debt-to-income ratio suppressed their yearly consumption by 0.5%. In 2009 the negative relationship starts to weaken and the point estimate is −0.028 in 2010:Q4, even though the debt-to-income ratio increased slightly between 2008:Q3 and 2010:Q2.

The results suggest that the negative effect of the debt-to-income ratio on consumption does not follow the business cycle. Fig. 3 showed that the average debt-to-income ratio increased most during 2009 and in 2009:Q4 an individual needed on average 134% of their yearly income to cover their total debt, compared to 124% of yearly income in 2008:Q4. Nevertheless, the estimated effect of the debt-to-income ratio decreases during the same period from −0.05 to −0.03, meaning that the negative effect on consumption weakens. This is consistent with the study by Dynan (2012), which shows that the negative relationship between leverage and consumption growth is stronger for 2005–2007 than for 2007–2009, although high standard errors make the estimates for 2005–2007 insignificant. Evidently it is important to compare the estimated coefficients for the recession period with the coefficients for the period of economic growth, otherwise the negative relationship between the debt-to-income ratio and consumption change during a recession might be interpreted as a phenomenon exclusive to a recession. The current estimations suggest that the debt-to-income ratio is holding back consumption growth but the effect appears not to be related to movements in the business cycle.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

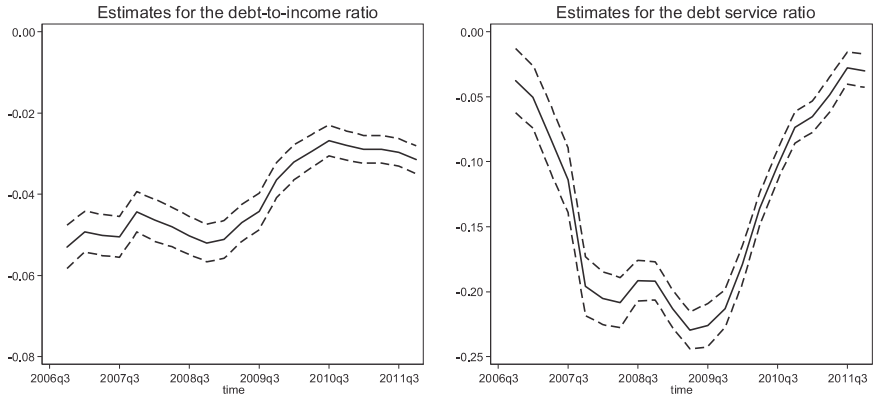


Fig. 5. The estimated quarterly coefficients of the lagged debt-to-income ratio and the lagged debt service ratio of Eq. (5) for each quarter from 2006:Q4 to 2011:Q4.

The relationship between the lagged debt service ratio and consumption change shows more variation over the business cycle. The coefficient is estimated to be quite modest at −0.038 in 2006:Q4. Given that the average debt service ratio is 0.25 in the same period, a yearly decline in consumption of only 1% can be associated with the debt service ratio during this period. In 2007 the estimated coefficient reaches −0.20 by 2007:Q4, meaning that the negative relationship between the debt service ratio and consumption change strengthens. The peak is in 2009:Q2 when the point estimate for the lagged debt service ratio is estimated to be −0.23. The negative relationship started to weaken in 2010 and the point estimate of the lagged debt service ratio reached the modest level of −0.03 by 2011:Q4.

The findings are similar to those of Dynan (2012), who estimates the coefficient for the debt service ratio to be negative in 2007–2009, meaning that debt obligations crowd out consumption in this period. However, the estimated coefficient in Dynan (2012) is close to zero in 2005–2007, implying no effect on consumption growth in this period.

There is clear evidence that the estimated coefficient for the debt-to-income ratio is relatively stable over the whole business cycle while the debt service ratio varies. The results suggest that the two debt variables do indeed reveal different channels through which household debt spills over into consumption. Apparently, the amplification effect is not induced by the effect of the fear of potential credit constraints as the estimated coefficient of the debt-to-income ratio does not follow the business cycle. The findings are consistent with the study of Meriküll (2015) which shows that it was credit demand that cut borrowing down during the recession rather than credit supply, suggesting that the fear of potential credit constraints in the future is not driving consumption decisions.

The results entail that the debt service ratio constrains consumption much more during the recession than during times of economic growth, all other things being equal. Evidently the results support the hypothesis that debt distress, which is measured by the debt service burden, is related to more conservative consumption behaviour. As explained in Section 2, in a recession households are usually more concerned about their financial situation and the concern may be stronger when households face additional compulsory expenses such as debt repayments. Higher compulsory expenses increase the precautionary motive, leading to more constrained consumption by households than if there are no debt repayments. Debt distress is probably low in a favourable economic environment when the risk of reverse shocks is low. This would explain the marginal negative relationship between the debt service ratio and consumption until 2007 and after 2010.

The increasing negative effect of the debt service ratio on consumption growth goes together with the increasing average debt service ratio shown in Fig. 4. The yearly debt payments increased substantially during 2007 compared to yearly income and reached their peak by the second half of 2009 when the debt service ratio was 0.36. In consequence, yearly consumption was suppressed by 8.3% on average in the sample during this period.

For comparison, aggregate real consumption declined by 15% in 2009:Q2. It is not possible to estimate the share of the drop in aggregate consumption induced by indebtedness, as current estimations give only the direct effect and it is more complicated to estimate the general equilibrium effects in the whole economy. Mian and Sufi (2012) show that the decrease in aggregate demand induced by the decrease in consumption leads to high unemployment, which has further implications for the economy. Nevertheless, the estimated drop in consumption growth due to indebtedness in the sample is substantial, suggesting an important role for debt in the aggregate consumption growth.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

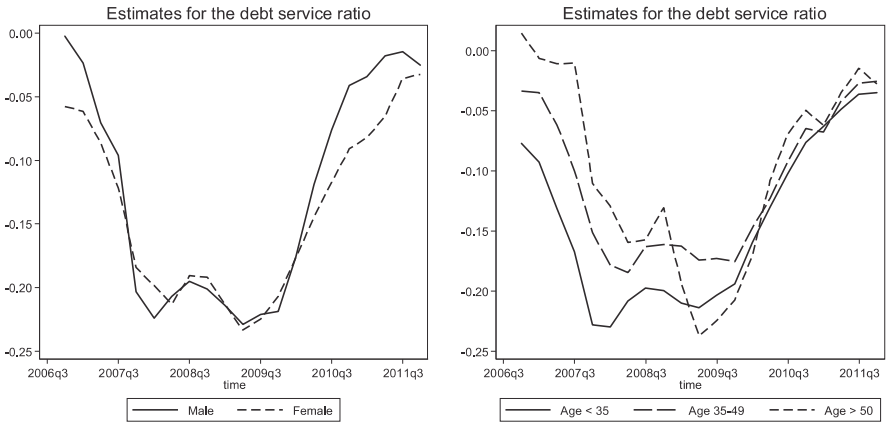


Fig. 6. The estimated quarterly coefficients of the lagged debt service ratio of Eq. (5) for different sub-samples from 2006:Q4 to 2011:Q4.

6.4. Robustness tests

As the current estimations were carried out for individuals and not households, additional robustness tests were used for different sub-samples. If spouses share their expenses unevenly, the effect of indebtedness on consumption might be different for males and females. This would complicate the interpretation of the results when individual-level data are used. The quarterly estimation results for the lagged debt service ratio are given in Fig. 6 and they show a very similar pattern for the estimations among females and males. The estimated coefficients of the lagged debt-to-income ratio are likewise very similar to the estimations given in Fig. 6 and so are not reported here. The similar estimations for the sub-samples suggest that the results are not affected when individual-level data are used rather than household level data.

Another set of estimations were run for different age groups. As younger age groups have higher income expectations than older age groups, the estimations for the sub-groups would reveal whether income expectations may influence the estimations.

The coefficient estimates vary slightly more at the pre-crisis level across age groups and the debt service ratio does not seem to suppress the consumption of individuals aged over 50, as the estimated coefficient is close to zero. However, the general trend in all age groups is similar, as the negative relationship between the lagged debt service ratio and consumption change is stronger during the recession and much weaker before and after the recession. This indicates there is a similar effect from the debt service ratio across different age groups, suggesting that the debt distress effect does not capture the effect of income expectations.

As an additional robustness test, estimations were carried out with a different income variable. The baseline estimations use inflow to the sight account from legal entities as an income variable. Alternatively, the total inflow to the sight account could be used as an income variable, although this variable is likely to be endogenous as individuals can arrange more inflows to their sight account if they plan to increase the outflow from the account, for example by making transactions between spouses or relatives or between different sight accounts. Another set of estimations of Eq. (5) was run where total inflow to the sight account was used as a proxy for earnings in the estimations of the debt-to-income ratio, the debt service ratio, the yearly total income and the asset-to-income ratio. The estimated quarterly coefficients for the lagged debt-to-income ratio are given in column (3) of Table C.1 in Appendix C and for the lagged debt service ratio in column (4). The point estimates are slightly lower for the debt-to-income ratio and slightly higher for the debt service ratio than the results of the baseline model given in columns (1) and (2). However, the dynamics of the estimated coefficients are very similar to the baseline model, indicating that the way the proxy for income is computed does not alter the main conclusions.

Finally, in columns (5) and (6) of Table C.1 in Appendix C the estimations excluding the main variable in the model, the income variable, are presented. The estimated coefficients for the debt-to-income ratio and the debt-service ratio are higher than in the baseline model in columns (1) and (2) in Table C.1. The differences in the estimations are induced by the omitted variable bias as explained in Section 3. Nevertheless, the coefficient for the debt-to-income ratio is quite stable over the sample period while the estimated coefficient for the debt service ratio varies over the business cycle. The strongest negative effect of the debt service ratio on consumption growth is seen in 2009:Q2 as in the baseline model. The robustness tests confirm that debt related variables capture different effects on consumption while the effect which is related to the debt service ratio follows the business cycle movements.

The upshot is that the comparison of the relationship between consumption and different debt variables reveals useful information about the possible mechanisms at work. The debt service ratio discloses more about the consumption decisions of households than the debt-to-income ratio does. The overall negative impact of indebtedness on consumption growth increases

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

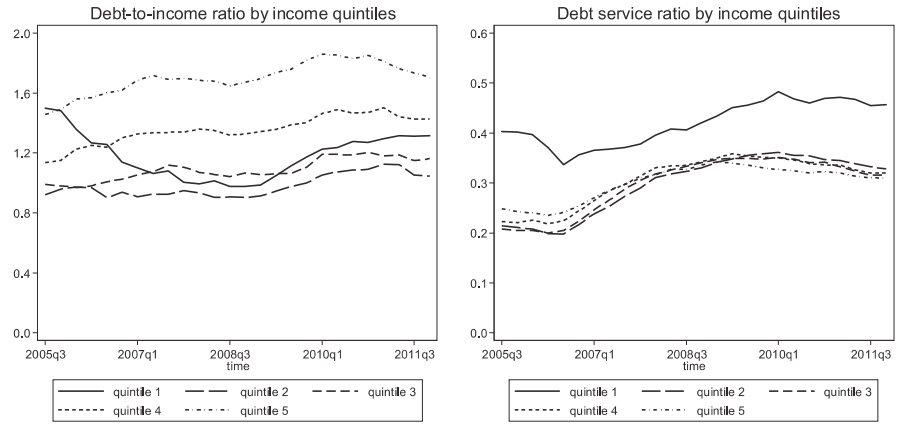


Fig. 7. The debt service ratio by income quintiles from 2005:Q3 to 2011:Q4.

substantially when the average debt service ratio increases and the effect is stronger during the recession. The dampening effect is marginal during the period of economic growth, meaning households only marginally decrease their consumption because of increasing debt burden levels.

The negative effect of the debt service ratio increases substantially during the recession, indicating that the debt service ratio has amplified the recession. The amplification of the recession occurs in two ways. First, households are more reluctant to consume when the debt service ratio is similar to what it was before the crisis. Second, the increase in the debt service ratio during the recession enhances the negative effect on consumption. The results indicate that debt servicing, which can be related to debt distress, played an important role in the consumption cutback during 2008–2009.

Although the results do not give a clear answer as to what exactly induces the increase in the debt distress in the recession, they still provide some policy suggestions. As the debt service ratio is important for consumption decisions, the outcome of policy measures which lower the debt service ratio, such as a decline in the interest rate, will spill over into the real economy by supporting household consumption. As mentioned in Section 5, housing loans contribute over 80% of the total loan volume in Estonia. As the majority of housing loans have been issued with an adjustable interest rate which is tied to Euribor, the interbank interest rate, changes in Euribor are quickly reflected in the debt repayment amounts. It has been calculated by a commercial bank that their mortgage customers have saved on average over 400 EUR per year from the declining interest rates (Rudzitis et al., 2012).

6.5. *The estimations for income quintiles*

Claessens and Perotti (2007) suggest that crises have strong redistributive consequences and Mian et al. (2013) highlight that debt causes the negative shocks to be distributed unevenly across households. They show heterogeneity in households' responses to negative shocks. Ampudia et al. (2014) suggest that the debt-to-income ratio for poor households increased during the recession and the heterogeneity in financial pressure might lead to different spending responses.

Further analysis of the different income quintiles would provide additional insights into the distribution of the negative relationship between indebtedness and consumption. Therefore it is investigated whether the impact of indebtedness on consumption differs across income groups. Individuals are divided into five income quintiles based on their yearly income and the mean debt-to-income ratio and debt service ratio are estimated for each income quintile and for each quarter.

Fig. 7 shows the evolution of the mean debt-to-income ratio and debt service ratio of individuals in different income quintiles. The average debt-to-income ratio is highest for individuals in the highest income quintile, indicating that individuals in higher income groups are more indebted than individuals in lower income groups. The average debt volume over the whole period is 98% of the yearly income for individuals in the second income quintile and 170% of the yearly income for individuals in the highest income quintile. In the 2000s, when financial deepening occurred in CEE countries, debts were mainly accumulated by higher-income households (Herzberg, 2010; Beck and Brown, 2011). The pattern is similar to that in the other euro-area countries where the Household Finance and Consumption Survey finds that the debt-to-income ratio is around 90% higher for households in the highest income quintile than for households in the second income quintile (European Central Bank (ECB), 2013, p. 65).

The rise in the debt-to-income ratio was steeper for individuals in higher income quintiles from 2006 to 2010, meaning that the debt volumes increased faster than income. The debt-to-income ratio increased by 30% for individuals in the fourth and fifth income quintiles while the increase was around 20% for individuals in the second and third income quintiles. The

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Table 3
Parameter estimates for the lagged debt-to-income ratio by income quintiles over years.

$Dtol_{t-4}$	2007	2008	2009	2010	2011
Quintile 1	−0.0783***	−0.0774***	−0.0652***	−0.0398***	−0.0358***
	(0.0085)	(0.0084)	(0.0072)	(0.0057)	(0.0044)
Quintile 2	−0.0864***	−0.0887***	−0.0735***	−0.0469***	−0.0410***
	(0.0055)	(0.0065)	(0.0051)	(0.0040)	(0.0038)
Quintile 3	−0.0775***	−0.0894***	−0.0767***	−0.0493***	−0.0457***
	(0.0042)	(0.0038)	(0.0035)	(0.0031)	(0.0030)
Quintile 4	−0.0741***	−0.0735***	−0.0726***	−0.0473***	−0.0501***
	(0.0033)	(0.0034)	(0.0032)	(0.0028)	(0.0025)
Quintile 5	−0.0665***	−0.0690***	−0.0637***	−0.0513***	−0.0560***
	(0.0035)	(0.0037)	(0.0050)	(0.0033)	(0.0030)
R^2	0.461				
No of groups	102 968				
No. of obs.	1 733 332				

Notes: FE estimation of Eq. (6). All explanatory variables and time dummies are included in the estimations but not reported. Standard errors are reported in parentheses below the coefficient estimates, SE estimates are robust to disturbances that are heteroskedastic and autocorrelated. ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

debt-to-income ratio in the lowest income quintile shows a different pattern as the ratio decreased significantly during the period of economic growth in 2006–2007 and increased again in 2010–2011; the dynamics of the ratio have been driven mainly by income changes. As individuals in different income quintiles experience different dynamics in the debt-to-income ratio, there might be differences in the impact of indebtedness on consumption.

The distribution of the debt service ratio across income groups is quite even, although individuals in the lowest income quintile have a somewhat higher debt service ratio than the individuals in other income groups, as the average debt service ratio for the lowest income quintile over all quarters is 0.42 while for the other income quintiles it is around 0.30. In the other euro-area countries the debt service ratio is similarly highest for households in the lowest income quintile, at the same level for households in the middle income groups, and lowest for households in the highest income quintile (ECB, 2013). The similar debt service ratio across different income groups may be related to the credit conditions of the credit supplier, which set the maximum share of income that can be used for debt servicing.

The increase in the debt service ratio occurs for individuals in all income groups in 2007–2009. Individuals in the second income quintile experience a rise in the debt service ratio of 80% between 2006:Q3 and 2010:Q1, while individuals in the highest income quintile face a rise of 38% during the same period.

In order to investigate the differences in the consumption change due to indebtedness between different income quintiles over the business cycle, the coefficients ϕ for the lagged debt-to-income ratio and ψ for the lagged debt service ratio in Eq. (4) are estimated for each income quintile separately while the income quintile of the previous year is used. *Yearly* estimations are implemented rather than quarterly estimations so that the estimations can be presented in a comprehensible form in a table. The model contains all the variables in Eq. (4). The following model was estimated:

$$\Delta_4 \log c_{it} = u_i + \sum_{p=1}^{5}\sum_{q=1}^{5} \beta_{pq} T_p \times I_q \times \log y_{it} + \sum_{p=1}^{5}\sum_{q=1}^{5} \phi_{pq} T_p \times I_q \times Dtol_{it-4} + \sum_{p=1}^{5}\sum_{q=1}^{5} \psi_{pq} T_p \times I_q \times Dsr_{it-4} + \sum_{p=1}^{5}\sum_{q=1}^{5} \alpha_{pq} T_p \times I_q \times Z'_{it-4} + X'_{it}\gamma + \tau_t + \varepsilon_{it} \quad (6)$$

The dummy T_p stands for the years from 2007 to 2011, the dummy I_q denotes the lagged income quintile q. The coefficients ϕ_{pq} and ψ_{pq} denote the impact of the lagged debt-to-income ratio and the lagged debt service ratio respectively on consumption for income quintile q in year p.

Table 3 gives the estimated coefficients of the lagged debt-to-income ratio in Eq. (6). The estimated coefficients are tabulated in a way that makes it easier to follow the evolution of the coefficient over the period in each quintile. The estimation years are given in columns and the income quintiles are given in rows. The estimations suggest that the negative relationship between the lagged debt-to-income ratio and consumption change is quite stable over the years for individuals in all income quintiles. The differences in the estimated coefficients for the different income groups and for the different years are statistically not significant, hence the estimations do not reveal any additional information about the relationship between the lagged debt-to-income ratio and consumption. The results confirm that the debt-to-income ratio does not amplify the recession in any of the income quintiles.

Table 4 provides the estimated coefficient for the lagged debt service ratio in Eq. (6). It shows that the negative relationship between the lagged debt service ratio and consumption growth strengthens in all income quintiles in 2007–2009 and weakens in 2010–2011. Hence the negative amplification effect is not concentrated among individuals with lower income but is spread across all income quintiles. The change in the estimated negative coefficient is the smallest for individuals in the highest income quintile,

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Table 4
Parameter estimates for the lagged debt service ratio by income quintiles over years.

Dsr_{t-4}	2007	2008	2009	2010	2011
Quintile 1	−0.0328	−0.0799***	−0.1350***	−0.0072	0.0700***
	(0.0378)	(0.0295)	(0.0237)	(0.0170)	(0.0206)
Quintile 2	−0.0386**	−0.1692***	−0.1818***	−0.0490***	0.0325**
	(0.0196)	(0.0144)	(0.0152)	(0.0093)	(0.0148)
Quintile 3	−0.0916***	−0.1940***	−0.2325***	−0.0757***	0.0294**
	(0.0157)	(0.0137)	(0.0133)	(0.0079)	(0.0124)
Quintile 4	−0.0475***	−0.1915***	−0.2228***	−0.0919***	−0.0267**
	(0.0148)	(0.0124)	(0.0125)	(0.0075)	(0.0113)
Quintile 5	−0.0905***	−0.1760***	−0.1387***	−0.0814***	−0.0365***
	(0.0162)	(0.0142)	(0.0150)	(0.0093)	(0.0125)
R^2	0.461				
No. of groups	102 968				
No. of obs.	1 733 332				

Notes: FE estimation of Eq. (6). All explanatory variables and time dummies are included in the estimations but not reported. Standard errors are reported in parentheses below the coefficient estimates, SE estimates are robust to disturbances that are heteroskedastic and autocorrelated. ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

where the estimated coefficient is 8.6 percentage points lower in 2008 than in 2007, while the decline of the estimated coefficient for individuals in the fourth income quintile is 14.4 percentage points. The individuals in the highest income quintile experience a smaller increase in the debt service ratio in 2007–2009 than do individuals in the lower income quintiles. This suggests that although indebtedness suppresses consumption during the recession in all income groups, the weakest amplification effect is observed in the highest income group.

To summarise, the negative effect of debt on consumption growth is not concentrated in any specific income group and indebtedness affects household consumption in a wide population group. No negative amplification effect on consumption growth is found for the debt-to-income ratio in any income quintiles, but the effect is apparent for the debt service ratio. It suggests that indebtedness plays an important role in determining the consumption decisions of *all* indebted households, regardless of their income level. The hypothesis that it is mainly low-income households that are affected by indebtedness rather than high-income households, and that they then cut their consumption does not appear to be valid in Estonia.

The reason for this may be related to the distribution of debt – Fig. 7 shows that high-income groups are the most indebted and the debt service ratio is at the same level as in other income groups, meaning that high-income households are as vulnerable to debt related shocks as low-income households are. This pattern apparently exists in all countries in which financial deepening mainly opened up borrowing opportunities for high-income groups, which have been less credit constrained than low-income households (Herzberg, 2010). The result applies not only to CEE countries but equally to other European countries, for example Andersen et al. (2014) find a similar negative relationship between leverage and consumption in all income groups in Denmark. The upshot of the findings is that the implications of indebtedness are spread among all indebted households, which makes the issue of alleviating the implications a highly relevant one.

7. Conclusions

This paper examines the role of household debt in amplifying the 2008–2009 recession. The topic is highly relevant for the CEE countries, in which household debt exhibited vigorous growth rates before the 2008–2009 crisis that was induced by the financial liberalisation in the first half of the 2000s. The paper uses a unique quarterly panel dataset from a financial institution covering the period from 2004:Q4 to 2011:Q4 and estimates the effect of indebtedness on consumption change over different parts of the business cycle.

The paper estimates a conventional consumption model augmented by two debt variables. The indebtedness is measured by two variables, the debt-to-yearly income ratio and the yearly debt service ratio. The first ratio indicates the debt management capacity while the second variable signifies the debt repayment burden. According to one hypothesis households may reduce their consumption due to the fear of potential credit constraints or because of precautionary concerns. It is argued that the debt-to-income ratio captures the effect of the fear of the potential credit constraints while the debt service ratio mainly indicates the debt distress effect. Additionally, there is a hypothesis that indebtedness amplifies the recession as households hold back consumption during a recession more than during times of economic growth.

Several broad patterns emerge from the results. Household indebtedness, measured as either the lagged debt-to-income ratio or the lagged debt service ratio, has a significant negative impact on household consumption growth over the whole period of 2006–2011. As both debt variables are negatively related with consumption growth, there are apparently several channels through which indebtedness affects consumption decisions. The dynamics of the effects of the two debt variables are different. The relationship between the debt-to-income ratio and consumption change is relatively stable, expressing a slightly decreasing relationship over the observed period which is apparently not related to the business cycle. The same pattern is observed among

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

individuals in all income quintiles, so no amplification of the recession is detected by the indebtedness expressed by the debt-to-income ratio.

On the other hand, the relationship between the lagged debt service ratio and consumption change varies over the business cycle. The debt service ratio is associated negatively with consumption growth during the recession in 2008–2009, while before and after the recession the negative relationship is marginal. Hence, an amplifying effect of the debt service ratio is found for the recession and the effect is slightly stronger for individuals in the lower income quintiles.

The results suggest that the debt service ratio contains useful information about the link between indebtedness and consumption. It is the debt service ratio rather than the debt-to-income ratio that captures the amplification of the recession. In the literature the importance of the measure of the debt-to-assets ratio is emphasised. The current results do not cancel out the collateral effect which is captured by the debt-to-assets ratio, as it is not possible to test the hypothesis with the current dataset. The current paper sheds light on additional debt related measures and implies that indebtedness constrains consumption through different channels.

The findings are valuable for policy suggestions as debt distress measured by the debt service ratio can be alleviated by lower interest rates, which was also observed in 2008–2009. As the majority of mortgages in Estonia have been issued as adjustable interest rate mortgages, a fall in the interest rate will affect most of the indebted households. Without this happening, the debt service ratio would probably have been higher and the negative effect on consumption might have been even stronger. Additionally, the evidence that the decline in consumption related to the debt service ratio is observed in all income groups highlights the importance of policy measures to relieve the debt payment burden, as this would affect most indebted households.

The general equilibrium effect of households' indebtedness on the aggregate economy is more complicated to estimate and the current study does not give any answer to this question. The drop in aggregate demand because of the drop in household consumption due to indebtedness will have further implications for the real economy through changes in the labour market, changes in prices or changes in saving behaviour. To some extent the broader implications of household debt for household behaviour and the real economy have been explored, see among others Mian and Sufi (2012) and Kukk (2014); but these issues are worth further investigation.

Acknowledgements

The author would like to thank Lennart Kitt for his help with the database, two anonymous referees and the editor of the Journal of Comparative Economics as well as seminar participants at TTÜ and Eesti Pank, participants at the HFCS network meeting in the ECB, the ECEE5 and EUA2015 conferences, and Alvar Kangur, Jaanika Meriküll, Atif Mian, Tiiu Paas, Giulia Rivolta, Karsten Staehr and Lenno Uusküla for useful comments. This work was supported by Estonian Target Financing grant no. SF0140059s12 and Base Financing grant no. B45/2015.

Appendix A

Tables A.1 and A.2

Table A.1
Panel summary statistics of the main variables.

Variable		Mean	St. dev.	Min	Max	Observations
Total spending						
	Overall	8.772	0.777	1.391	14.466	$N = 2\,597\,130$
	Between		0.61	5.107	11.944	$n = 107\,859$
	Within		0.439	2.181	13.683	$T = 24$
Total income						
	Overall	8.737	0.728	5.177	15.289	$N = 2\,597\,930$
	Between		0.655	6.322	11.972	$n = 107\,859$
	Within		0.372	4.840	14.099	$T = 24$
Total income from legal entities						
	Overall	8.530	0.773	−3.026	14.094	$N = 2\,597\,930$
	Between		0.706	2.327	11.636	$n = 107\,859$
	Within		0.404	−0.074	13.341	$T = 24$
Debt-to-income ratio						
	Overall	0.574	1.647	0	21.079	$N = 2\,597\,930$
	Between		1.527	0	17.497	$n = 107\,859$
	Within		0.824	−14.385	19.682	$T = 24$
Debt service ratio						
	Overall	0.138	0.265	0	2.219	$N = 2\,597\,930$
	Between		0.221	0	1.969	$n = 107\,859$
	Within		0.159	−1.496	2.252	$T = 24$

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Table A.2
Summary statistics of the debt-to-income ratio and the debt service ratio by quarter, conditional on ownership of debt.

		Debt-to-income ratio			Debt service ratio		
	Quarter	Mean	Median	St. dev.	Mean	Median	St. dev.
2005	3	1.184	0.213	2.106	0.234	0.166	0.233
2005	4	1.199	0.215	2.107	0.231	0.165	0.230
2006	1	1.232	0.219	2.175	0.232	0.165	0.232
2006	2	1.235	0.235	2.176	0.225	0.160	0.228
2006	3	1.229	0.256	2.168	0.227	0.163	0.227
2006	4	1.248	0.295	2.197	0.245	0.173	0.242
2007	1	1.268	0.301	2.240	0.265	0.184	0.267
2007	2	1.280	0.320	2.248	0.282	0.190	0.292
2007	3	1.282	0.329	2.238	0.297	0.198	0.310
2007	4	1.277	0.334	2.220	0.312	0.205	0.325
2008	1	1.266	0.336	2.188	0.326	0.212	0.337
2008	2	1.253	0.329	2.171	0.335	0.216	0.345
2008	3	1.230	0.320	2.118	0.337	0.219	0.345
2008	4	1.242	0.313	2.146	0.346	0.225	0.349
2009	1	1.254	0.307	2.169	0.354	0.232	0.353
2009	2	1.281	0.305	2.218	0.360	0.236	0.355
2009	3	1.308	0.296	2.287	0.359	0.237	0.355
2009	4	1.347	0.297	2.342	0.359	0.239	0.352
2010	1	1.404	0.299	2.447	0.362	0.241	0.355
2010	2	1.417	0.291	2.478	0.357	0.238	0.351
2010	3	1.413	0.284	2.486	0.351	0.233	0.348
2010	4	1.427	0.287	2.477	0.351	0.233	0.349
2011	1	1.430	0.280	2.495	0.348	0.229	0.347
2011	2	1.405	0.270	2.462	0.340	0.223	0.343
2011	3	1.375	0.266	2.408	0.333	0.217	0.338
2011	4	1.370	0.259	2.410	0.333	0.215	0.340
Total		1.303	0.288	2.281	0.314	0.207	0.323

Appendix B

Table B.1–Table B.4

Table B.1
Definitions of all the variables used in the empirical model with summary statistics.

Variable	Definition	Mean	St. dev.
$\log y_{it}$	Logarithm of real yearly inflow from legal entities to sight accounts of an individual i in quarter t, in EUR in 2005 prices	8.530	0.773
$\log c_{it}$	Logarithm of real yearly outflows of sight account of an individual i in quarter t, excluding transactions between saving and investment accounts, in EUR in 2005 prices	8.772	0.777
$DtoI_{it}$	Debt-to-yearly income ratio; debt stock is measured at the end of quarter t and income is the sum of the income of the four previous quarters	0.574	1.647
Dsr_{it}	The ratio of annual debt service payments to annual income in quarter t	0.138	0.265
Finasset_{it}	Ratio of financial assets to yearly income from legal entities at the end of quarter t. Financial assets include deposits, investment funds, stocks, bonds and pension funds	0.660	51.923
Overdue_{it}	Dummy = 1 if individual has had any repayment problems to the commercial bank in quarter t, otherwise = 0	0.008	0.091
Lifeins_{it}	Dummy = 1 if the individual has a life insurance contract in quarter t, otherwise = 0	0.056	0.230
Pensionlife_{it}	Dummy = 1 if the individual has a pension insurance contract in quarter t, otherwise = 0	0.090	0.287
NewH_{it}	Dummy = 1 if the individual owns a housing loan in quarter t while not having any housing loan in quarter $t-1$, otherwise = 0	0.002	0.047
AddH_{it}	Dummy = 1 if the size of the individual's housing loan in quarter t exceeds the housing loan in quarter $t-1$ and the individual owns the housing loan in quarter $t-1$, otherwise = 0	0.005	0.072

Table B.2
Unit root tests.

	(1) Harris–Tzavalis test Statistic (ρ)	(2) Im–Pesaran–Shu test Statistic ($Z_{\bar{t}-\text{bar}}$)
$\Delta_4 \log c_{it}$	0.6906***	−114.4***
$\Delta_4 \log y_{it}$	0.7170***	−100.6***
$DtoI_{it}$	0.7992***	−15.44***
Dsr_{it}	0.7993***	−254.9***
Finasset_{it}	0.7367***	−27.35***

Notes: Unit root tests with panel means included, time trend not included, and cross-sectional means removed. Harris–Tzavalis test assumes a common autoregressive parameter and the null hypothesis is that panels contain unit roots while the alternative hypothesis is that panels are stationary. Im–Pesaran–Shu test assumes panel-specific autoregressive parameters and the null hypothesis is that all panels contain unit roots while the alternative hypothesis is that some panels are stationary. ***, ** and * indicate that the null hypothesis is rejected at the 1%, 5% and 10% level respectively.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Table B.3
Estimations for the full sample period 2006:Q4–2011:Q4.

	(1)	(2)	(3)	(4)	(5)
$\Delta_4 \log y_t$	0.5460***	0.5447***	0.5491***	0.5749***	0.5499***
	(0.0035)	(0.0035)	(0.0035)	(0.0030)	(0.0024)
$Dtoi_{t-4}$	−0.0478***	…	−0.0367***	−0.0354***	−0.0022***
	(0.0016)		(0.0016)	(0.0016)	(0.0004)
Dsr_{t-4}	…	−0.1792***	−0.1397***	−0.1872***	−0.0644***
		(0.0038)	(0.0040)	(0.0039)	(0.0022)
Finasset_{t-4}	0.0220*	0.0217*	0.0218*	0.0198*	0.0079**
	(0.0132)	(0.0130)	(0.0131)	(0.0120)	(0.0033)
Overdue_{t-4}	−0.0421***	−0.0354***	−0.0334***	−0.0379***	−0.0550***
	(0.0046)	(0.0046)	(0.0046)	(0.0046)	(0.0042)
$\Delta_4 \text{ lifeins}_t$	0.0494***	0.0490***	0.0483***	0.0627***	0.0518***
	(0.0052)	(0.0052)	(0.0052)	(0.0053)	(0.0045)
$\Delta_4 \text{ pensionlife}_t$	0.0369***	0.0346***	0.0344***	0.0519***	0.0354***
	(0.0045)	(0.0045)	(0.0045)	(0.0045)	(0.0039)
$\sum_{s=0}^{8} newH_{t-s}$	Yes	Yes	Yes	Yes	Yes
$\sum_{s=0}^{8} addH_{t-s}$	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	No	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	No
R^2	0.2803	0.2804	0.2819	0.2761	0.2931
No. of groups	102 968	102 968	102 968	102 968	102 968
No. of obs.	1 733 332	1 733 332	1 733 332	1 733 332	1 733 332

Notes: FE estimation of Eq. (4). Standard errors are reported in parentheses below the coefficient estimates, SE estimates are robust to disturbances that are heteroskedastic and autocorrelated. ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

Table B.4
Robustness of the estimations for the full sample period 2006:Q4–2011:Q4.

	(1)	(2)	(3)	(4)	(5)
$\Delta_4 \log y_t$	…	0.5536***	0.5491***	0.5490***	0.7481***
		(0.0026)	(0.0035)	(0.0035)	(0.0036)
$Dtoi_{t-4}$	−0.0019	−0.0363***	−0.0367***	−0.0367***	−0.0586***
	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0022)
Dsr_{t-4}	−0.0533***	−0.1411***	−0.1406***	−0.1403***	−0.0999***
	(0.0047)	(0.0039)	(0.0040)	(0.0040)	(0.0046)
Finasset_{t-4}	0.0365*	…	0.0218*	0.0218*	0.1393***
	(0.0199)		(0.0131)	(0.0131)	(0.0360)
Overdue_{t-4}	−0.0381***	−0.0336***	…	−0.0337***	−0.0234***
	(0.0058)	(0.0046)		(0.0046)	(0.0036)
$\Delta_4 \text{ lifeins}_t$	0.0490***	0.0484***	0.0484***	…	0.0425***
	(0.0059)	(0.0052)	(0.0052)		(0.0045)
$\Delta_4 \text{ pensionlife}_t$	0.0298***	0.0344***	0.0345***	…	0.0292***
	(0.0052)	(0.0045)	(0.0045)		(0.0038)
$\sum_{s=0}^{8} newH_{t-s}$	Yes	Yes	Yes	Yes	No
$\sum_{s=0}^{8} addH_{t-s}$	Yes	Yes	Yes	Yes	No
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.108	0.2798	0.2819	0.2817	0.4383
No. of groups	102 968	102 968	102 968	102 968	102 968
No. of obs.	1 733 332	1 733 332	1 733 332	1 733 332	1 733 332

Notes: FE estimation of Eq. (4). Standard errors are reported in parentheses below the coefficient estimates, SE estimates are robust to disturbances that are heteroskedastic and autocorrelated. ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Appendix C

Table C.1

Table C.1
Estimated quarterly parameters for the debt-to-income ratio and the debt service ratio using different yearly income measures.

Income measure		(1) Income from legal entities		(3) Total inflow to sight account		(5) W/o income	
			(2)		(4)		(6)
Year	Q	Debt-to-income ratio	Debt service ratio	Debt-to-income ratio	Debt service ratio	Debt-to-income ratio	Debt service ratio
2006	4	−0.0530***	−0.0376***	−0.0780***	0.0178	−0.0281***	0.1291***
		(0.0027)	(0.0126)	(0.0031)	(0.0133)	(0.0035)	(0.0173)
2007	1	−0.0492***	−0.0503***	−0.0724***	−0.0012	−0.0226***	0.1308***
		(0.0026)	(0.0122)	(0.0029)	(0.0128)	(0.0033)	(0.0166)
2007	2	−0.0501***	−0.0821***	−0.0758***	−0.0235*	−0.0247***	0.1386***
		(0.0026)	(0.0133)	(0.0029)	(0.0132)	(0.0033)	(0.0173)
2007	3	−0.0505***	−0.1141***	−0.0757***	−0.0698***	−0.0233***	0.1017***
		(0.0026)	(0.0127)	(0.0031)	(0.0126)	(0.0034)	(0.0163)
2007	4	−0.0444***	−0.1958***	−0.0706***	−0.1659***	−0.0167***	0.0129
		(0.0025)	(0.0116)	(0.0030)	(0.0126)	(0.0034)	(0.0160)
2008	1	−0.0464***	−0.2051***	−0.0733***	−0.1854***	−0.0165***	−0.0171
		(0.0026)	(0.0103)	(0.0033)	(0.0117)	(0.0037)	(0.0145)
2008	2	−0.0480***	−0.2084***	−0.0751***	−0.1986***	−0.0152***	−0.0556***
		(0.0025)	(0.0098)	(0.0032)	(0.0145)	(0.0036)	(0.0184)
2008	3	−0.0502***	−0.1915***	−0.0780***	−0.1771***	−0.0142***	−0.0587***
		(0.0024)	(0.0081)	(0.0030)	(0.0129)	(0.0034)	(0.0168)
2008	4	−0.0520***	−0.1918***	−0.0801***	−0.1863***	−0.0150***	−0.0774***
		(0.0023)	(0.0074)	(0.0029)	(0.0114)	(0.0033)	(0.0145)
2009	1	−0.0511***	−0.2132***	−0.0781***	−0.2150***	−0.0104***	−0.1226***
		(0.0024)	(0.0074)	(0.0030)	(0.0128)	(0.0033)	(0.0162)
2009	2	−0.0470***	−0.2297***	−0.0734***	−0.2313***	−0.0033	−0.1440***
		(0.0023)	(0.0073)	(0.0029)	(0.0138)	(0.0033)	(0.0175)
2009	3	−0.0442***	−0.2260***	−0.0704***	−0.1948***	−0.0006	−0.1096***
		(0.0023)	(0.0085)	(0.0028)	(0.0096)	(0.0032)	(0.0126)
2009	4	−0.0365***	−0.2131***	−0.0605***	−0.1539***	0.0105***	−0.0805***
		(0.0022)	(0.0074)	(0.0028)	(0.0121)	(0.0032)	(0.0150)
2010	1	−0.0321***	−0.1791***	−0.0530***	−0.1191***	0.0143***	−0.0465***
		(0.0022)	(0.0074)	(0.0028)	(0.0086)	(0.0031)	(0.0113)
2010	2	−0.0294***	−0.1360***	−0.0489***	−0.0999***	0.0172***	−0.0153*
		(0.0021)	(0.0062)	(0.0026)	(0.0067)	(0.0030)	(0.0091)
2010	3	−0.0268***	−0.1030***	−0.0461***	−0.0614***	0.0190***	0.0287***
		(0.0019)	(0.0060)	(0.0025)	(0.0061)	(0.0028)	(0.0085)
2010	4	−0.0280***	−0.0737***	−0.0464***	−0.0389***	0.0166***	0.0668***
		(0.0019)	(0.0062)	(0.0023)	(0.0091)	(0.0027)	(0.0119)
2011	1	−0.0290***	−0.0653***	−0.0481***	−0.0331***	0.0137***	0.0826***
		(0.0018)	(0.0061)	(0.0022)	(0.0095)	(0.0026)	(0.0123)
2011	2	−0.0290***	−0.0485***	−0.0483***	−0.0020	0.0137***	0.1241***
		(0.0017)	(0.0069)	(0.0021)	(0.0087)	(0.0024)	(0.0116)
2011	3	−0.0297***	−0.0279***	−0.0489***	0.0158*	0.0139***	0.1455***
		(0.0017)	(0.0063)	(0.0022)	(0.0085)	(0.0025)	(0.0115)
2011	4	−0.0315***	−0.0300***	−0.0512***	0.0164**	0.0123***	0.1518***
		(0.0018)	(0.0065)	(0.0022)	(0.0081)	(0.0025)	(0.0109)
R^2	0.288	0.466	0.141				
No. of groups	102 968	102 968	102 968				
No. of obs.	1 733 332	1 733 332	1 733 332				

Notes: FE estimation of Eq. (5). All explanatory variables and time dummies are included in the estimations but not shown in the table. Standard errors are reported in parentheses below the coefficient estimates, SE estimates are robust to disturbances that are heteroskedastic and autocorrelated. ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

References

Aben, Madis, Kukk, Merike, Staehr, Karsten, 2012. Housing equity withdrawal and consumption dynamics in Estonia 2002-2011. Research in Economics and Business: Central and Eastern Europe 4 (1), 19–40.
Ampudia, Miguel, Akmaral Pavlickova, Jiri Slacalek, and Edgar Vogel, 2014. Household heterogeneity in the euro area since the onset of the Great Recession. European Central Bank Working Paper no. 1705.
Andersen, Asger L., Charlotte Duus, and Thais L. Jensen, 2014. Household debt and consumption during the crisis: evidence from Danish micro data. Danmarks Nationalbank, Working Paper no. 89.
Attanasio, Orazio P., Blow, Laura, Hamilton, Robert, Leicester, Andrew, 2009. Booms and busts: consumption, house prices and expectations. Economica 76 (301), 20–50.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Attanasio, Orazio P., Weber, Guglielmo, 2010. Consumption and saving: models of intertemporal allocation and their implications for public policy. Journal of Economic Literature 48 (3), 693–751.
Baker, Scott R., 2014. Debt and the consumption response to household income shocks. Stanford University, Unpublished manuscript available at http://web.stanford.edu/~srbaker/Papers/Baker_DebtConsumption.pdf.
Barba, Aldo, Pivetti, Massimo, 2009. Rising household debt: its causes and macroeconomic implications—a long-period analysis. Cambridge Journal of Economics 33 (1), 113–137.
Barrell, Ray, E. Philip Davis, Tatiana Fic, and Ali Orazgani, 2009. Household debt and foreign currency borrowing in new member states of the EU. Brunel University Economics and Finance Working Paper no. 09-23.
Bassanetti, Antonio, and Francesco Zollino, 2010. The effects of housing and financial wealth on personal consumption: aggregate evidence for Italian households. In: De Bandt, O., Knetch, T., Penalosa, J., and Zollino, F. (Eds.), Housing Markets in Europe. Springer, Berlin, Heidelberg, pp. 307-336.
Beck, Thorsten, Brown, Martin, 2011. Use of Banking Services in Emerging Markets: Household-Level Evidence. Centre for Economic Policy Research.
Brixiova, Zuzana, Vartia, Laura, Wörgötter, Andreas, 2010. Capital flows and the boom–bust cycle: the case of Estonia. Economic Systems 34 (1), 55–72.
Brown, Martin, Ralph De Haas, and Pauline Grosjean, 2012. Household debt and economic vulnerability during the Great Recession. Mimeo, the European Bank for Reconstruction and Development.
Brown, Sarah, Taylor, Karl, Price, StephenWheatley, 2005. Debt and distress: evaluating the psychological cost of credit. Journal of Economic Psychology 26 (5), 642–663.
Buiter, Willem H, 2010. Housing wealth isn't wealth. Economics E-Journal 4 (22), 1–29 http://www.economics-ejournal.org/economics/journalarticles/2010-22 .
Cecchetti, Stephen G., Madhusudan S. Mohanty, and Fabrizio Zampolli, 2011. The real effects of debt. Bank for International Settlements, Working Paper no. 352.
Chien, Yi-Wen, Devaney, SharonA., 2001. The effects of credit attitude and socioeconomic factors on credit card and installment debt. Journal of Consumer Affairs 35 (1), 162–179.
Claessens, Stijn, Perotti, Enrico, 2007. Finance and inequality: channels and evidence. Journal of Comparative Economics 35 (4), 748–773.
Cooper, Daniel, 2012. U.S. Household Deleveraging: What Do the Aggregate and Household-Level Data Tell Us? Federal Reserve Bank of Boston, Public Policy Briefs 12-2.
Coricelli, Fabrizio, Fabio Mucci, and Debora Revoltella, 2006. Household credit in the New Europe: lending boom or sustainable growth? CEPR Discussion Paper no. 5520.
Debelle, Guy, 2004. Macroeconomic implications of rising household debt. Bank for International Settlements, Working Paper no. 153.
Demirgüç-Kunt, Asli, and Leora F. Klapper, 2012. Measuring financial inclusion: the Global Findex Database. The World Bank Policy Research, Working Paper no. 6025
Dynan, Karen, 2012. Is a household debt overhang holding back consumption? Brookings Papers on Economic Activity 43 (Spring), 299–362.
Dynan, Karen, Edelberg, Wendy, 2013. The relationship between leverage and household spending behavior: evidence from the 2007-2009 survey of consumer finances. Federal Reserve Bank of St. Louis Review 95, 425–448 (September/October).
European Central Bank (ECB), 2013. The Eurosystem Household Finance and Consumption Survey. Results from the First Wave Statistics Paper Series No. 2.
Eurobarometer, 2014. Coping with the Crisis. Aggregate Report. Eurobarometer Qualitative study, the European Commission, November. Available at http://ec.europa.eu/public_opinion/archives/quali/ql_6342_crisis_en.pdf (doi:10.2767/55315).
Eggertsson, Gauti B, Krugman, Paul, 2012. Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. The Quarterly Journal of Economics doi:10.1093/qje/qjs023.
Friedman, Milton, 1957. A Theory of the Consumption Function. Princeton University Press, Princeton and Oxford.
Girouard, Nathalie, Mike Kennedy, and Christophe Andre, 2006. Has the rise in debt made households more vulnerable? OECD Working Paper no. 535.
Glick, Reuven, Lansing, KevinJ., 2010. Global Household Leverage, House Prices, and Consumption. Federal Reserve Bank of San Francisco Economic Letter No. 01.
Guerrieri, Luca, and Matteo Iacoviello, 2013. Collateral constraints and macroeconomic asymmetries. Federal Reserve Board, International Finance Discussion Paper no. 1082.
Harris, Richard DF, Tzavalis, Elias, 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics 91 (2), 201–226.
Herzberg, Valerie, 2010. Assessing the risk of private sector debt overhang in the Baltic countries. International Monetary Fund, Working Paper no. 250.
Hilbers, Paul, Inci Otker-Robe, Ceyla Pazarbasioglu, and Gudrun Johnsen, 2005. Assessing and managing rapid credit growth and the role of supervisory and prudential policies. IMF Working Paper no. 151.
Im, Kyung So, Hashem Pesaran, M., Shin, Yongcheol, 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115 (1), 53–74.
Jappelli, Tullio, Pistaferri, Luigi, 2010. The consumption response to income changes. Annual Review of Economics 2 (1), 479–506.
Jordà, Òscar, Moritz H.P. Schularick, and Alan M. Taylor, 2011. When credit bites back: leverage, business cycles, and crises. National Bureau of Economic Research, Working Paper no. 17621.
Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti, 2013. Household leveraging and deleveraging. NBER Working Paper no. 18941.
King, Mervin, 1994. Debt deflation: theory and evidence. European Economic Review 38 (2), 419–445.
Kukk, Merike, 2014. Distinguishing the components of household financial wealth: the impact of liabilities on assets in Euro Area countries. Bank of Estonia Working Paper no. 2.
Kukk, Merike, Kulikov, Dmitry, Staehr, Karsten, 2015. Estimating consumption responses to income shocks of different persistence using self-reported income measures. Review of Income and Wealth doi:10.1111/roiw.12163, forthcoming.
Kukk, Merike, Staehr, Karsten, 2014. Income underreporting by households with business income. Evidence from Estonia. Post-Communist Economies 26 (2), 257–276.
Martinez-Carrascal, Carmen, and Ana del Rio, 2004. Household borrowing and consumption in Spain: a VECM approach. Bank of Spain Working Paper no. 0421.
Meriküll, Jaanika, 2015. Households borrowing during a creditless recovery. Emerging Markets Finance and Trade doi:10.1080/1540496X.2015.1048154, forthcoming.
Mian, Atif, Sufi, Amir, 2010a. The great recession: Lessons from microeconomic data. American Economic Review 100 (2), 51–56.
Mian, Atif, Sufi, Amir, 2010b. Household leverage and the recession of 2007–09. IMF Economic Review 58 (1), 74–117.
Mian, Atif, Sufi, Amir, 2011. House prices, home equity-based borrowing, and the US household leverage crisis. American Economic Review 101, 2132–2156.
Mian, Atif, and Amir Sufi, 2012. What explains high unemployment? The aggregate demand channel. National Bureau of Economic Research. Working Paper no. 17830.
Mian, Atif, Rao, Kamalesh, Sufi, Amir, 2013. Household balance sheets, consumption, and the economic slump. Quarterly Journal of Economics 128 (4), 1687–1726.
Modigliani, Franco, Brumberg, Richard, 1954. Utility analysis and the consumption function: an interpretation of cross-section data. In: Kurihara, Kenneth K. (Ed.), Post Keynesian Economics. Rutgers University Press, New Brunswick, pp. 388–436.
Morley, James C., 2007. The slow adjustment of aggregate consumption to permanent income. *Journal of Money, Credit and Bank*ing 39 (2–3), 615–638.
Rudzitis, Edmunds, Julita Varanauskiene, Triin Messimas, and Merike Kukk, 2012. Baltic Household Outlook, October. SEB Group. Available at http://www.seb.ee/sites/default/files/web/files/uudised/baltic_household_outlook_oct_2012.pdf.

Please cite this article as: M. Kukk, How did household indebtedness hamper consumption during the recession? Evidence from micro data, Journal of Comparative Economics (2015), http://dx.doi.org/10.1016/j.jce.2015.07.004

Appendix 3. Publication III

Reprint from:

Kukk, Merike, and Karsten Staehr (2014). Income underreporting by households with business income. Evidence from Estonia. *Post-Communist Economies*, vol. 26, no. 2, pp. 257-276. Available online at: www.tandfonline.com/loi/cpce20, DOI: 10.1080/14631377.2014.904110

Post-Communist Economies, 2014
Vol. 26, No. 2, 257–276, http://dx.doi.org/10.1080/14631377.2014.904110

Routledge
Taylor & Francis Group

Income underreporting by households with business income: evidence from Estonia

Merike Kukk[a]* and Karsten Staehr[a,b]

[a]*Tallinn University of Technology, Tallinn, Estonia;* [b]*Eesti Pank, Tallinn, Estonia*

(*Received 9 July 2013; final version received 5 November 2013*)

This article estimates the extent of income underreporting by households with business income relative to households of wage earners in Estonia. It uses a modified version of the methodology pioneered by Pissarides and Weber. The extent of income underreporting is estimated by comparing food Engel curves for households with and without reported business income. The baseline result is that the reported total income of households with business income above 20% of total income must be multiplied by 2.6 in order to attain the same propensity to food consumption as households of wage earners. In this sense, households with business income underreport 62% of their 'true' total income. Households with reported business income above 0 but below 20% also underreport income but to a lesser extent. The estimates are higher than those found for developed countries but consistent with other studies of unreported activities in transition countries.

The shadow economy encapsulates income or production not reported to the authorities. The aim of such underreporting is typically to evade taxation but may also be to hide an illegal activity or to avoid an administrative burden. Many studies focus on estimates of the *aggregate* size of the shadow economy; see Schneider *et al.* (2010) for a review of various methods used to provide such estimates and some results.

It is equally important for informed policy making to obtain estimates of the extent of specific forms of underreporting. Different income components such as wage payments, income from self-employment or corporate income may be subject to underreporting to the tax board and other authorities (Black *et al.* 2012). Reliable estimates of the size of the shadow economy are important for assessing the distribution of income in society and for analysis of the incidence and welfare effects of taxation policies.

Underreporting by individuals who earn their income from self-employment or other business-related activities has received limited attention in the empirical literature. This is unfortunate since evasion of the tax on business income appears to be widespread. Slemrod (2007) uses data in a report from the US Internal Revenue Service and estimates that while only 1% of wage and salary income and 12% of capital income were left unreported in 2001, as much as 43% of the business income of individuals was not reported to the tax authorities in the USA. There are no similar analyses of the contribution of different forms of tax evasion in other countries but it is reasonable to assume a similar pattern elsewhere. As the business income of individuals is largely the subject of individual reporting, there is generally ample scope for evading taxes on business income. Some studies even argue that the chance to engage in tax evasion could be a reason for people to choose self-employment or engagement in other business-related activities (Bruce 2000).

*Corresponding author. Email: merike.kukk@ttu.ee

It is challenging to estimate the prevalence of underreporting of business income, as the main reason for underreporting is to avoid information on the true income being made available to the authorities. Unlike most employed individuals, individuals with business-related income have substantial discretion about the information provided to the tax authorities. If individuals underreport their income to the tax authorities, they may also underreport their income to other collectors of data. This point was illustrated by the 2007 Eurobarometer survey on undeclared work in which the data on underreporting by the self-employed were deemed unreliable because the data in many cases were inconsistent with other studies (Eurobarometer 2007).

Pissarides and Weber (1989) introduced an innovative methodology for providing estimates of income underreporting by the self-employed, using data from household budget surveys on, *inter alia*, income and consumption. The starting point is the Engel curve to food consumption, which posits that food consumption and income are closely related for otherwise comparable households. Different propensities to food consumption between the self-employed and wage earners may consequently stem from underreporting of income. The 'true' income, comprising both reported and unreported income, which would match the level of food consumption can then be computed.

Although individuals do not have any direct incentive to underreport their business-related income to a household budget survey, they may still feel compelled to provide consistent data about their income to all data collectors, particularly if they have any suspicion that the information will be shared with the tax authorities. Individuals will not have a similar incentive to underreport consumption levels as there is no direct tax evasion involved. Moreover, household budget surveys typically require the respondents to provide detailed information on their purchases of individual items and this can lead to a fairly precise estimation of their consumption levels. In short, it is reasonable to assume that household budget surveys contain relatively reliable consumption data, while data on business income may be subject to underreporting (Hurst *et al.* 2014).

This study uses a modified version of the methodology pioneered by Pissarides and Weber (1989) to provide estimates of income underreporting by households with business income relative to households of wage earners in Estonia. The analysis is based on data from the Estonian Household Budget Survey (HBS) for 2002–07, a period when Estonia experienced rapid economic growth. The method has been used in a handful of studies of tax evasion by the self-employed but mainly for developed countries.

This article contributes to the literature in four ways. First, we focus on Estonia, a transition country from Central and Eastern Europe. Like other transition countries, Estonia differs from developed countries by having a lower income level and evolving institutional and administrative systems. The shadow economy is generally larger in the transition countries in Central and Eastern Europe than in West European countries (Tafenau *et al.* 2010, Schneider *et al.* 2010). Studies of tax evasion in transition economies have hitherto focused mainly on tax morale (Torgler 2003) or informal employment and 'envelope wages', i.e. undeclared cash payments for salaries (see e.g. Williams 2009, Merikül and Staehr 2010, Schneider 2011). Other forms of tax evasion, including income underreporting by individuals with business income, have received less attention. Kim *et al.* (2009) use the methodology of Pissarides and Weber (1989) to assess income underreporting by self-employed households in Russia in the 1990s but the methodology has not previously been used to assess underreporting by the self-employed for any of the transition countries of Central and Eastern Europe.[1]

Second, Estonia joined the European Union in 2004 along with a number of other European transition countries. In the run-up to accession many reforms were passed in

order to achieve compliance with the *acquis communautaire*. The result was numerous changes in company law, taxation rules, industrial policy and government institutions. Meanwhile the Estonian economy grew very rapidly, in part due to the improved confidence stemming from the accession process. This article seeks to analyse whether these changes have affected income underreporting by households with business income.

Third, the study follows the methodology of Pissarides and Weber (1989) and identifies households with a high probability of underreporting based on the share of reported business income in total reported household income. The latter definition assumes a given threshold and we investigate the importance of the choice of threshold value.

Fourth, the methodology of Pissarides and Weber (1989) requires data on the permanent income of households. The income is usually not observed and additional assumptions are required to compute a proxy, and this usually gives rise to a range of estimates of the underreporting by self-employed households, with a lower and an upper bound. The Estonian Household Budget Survey makes it possible to disentangle transitory and permanent income using self-reported information, and we are thus able to provide a point estimate of the extent of underreporting. Kim *et al.* (2009) also provide a point estimate but use a statistical filtering methodology to extract a measure of permanent income.

The rest of the article is organised as follows: first we provide a review of the empirical literature that uses food Engel curves to estimate income underreporting of the self-employed or households with business income. Next we discuss the methodology of Pissarides and Weber (1989) and develop the methodology used in the study. Then we present the data from the Estonian Household Budget Survey used in the empirical analysis and examine the main properties of the consumption and income data. The results of the estimations are then presented and the final section summarises the empirical findings.

Literature estimating income underreporting by the self-employed

Pissarides and Weber (1989), henceforth P&W, introduced an innovative methodology for providing estimates of income underreporting by self-employed households. They consider a household to be self-employed when business-related income exceeds a given share of its reported income. They calculate the unreported taxable income of the self-employed in the UK using income and expenditure data from the Family Expenditure Survey (FES). A detailed description of the methodology is provided below. They find that to obtain the 'true' income, i.e. the sum of both reported and unreported income, the reported total income must be multiplied by a factor of 1.51–1.64 for the blue-collar self-employed and by 1.28–1.54 for the white-collar self-employed. In other words, the blue-collar self-employed households leave 34–39% of their 'true' total income unreported, while white-collar self-employed households leave 23–35% unreported.

Estimations for the USA using the P&W methodology have been done by Hurst *et al.* (2014). Using data from the Consumer Expenditure Survey (1980–2003) and the Panel Study of Income Dynamics (1980–1997), they find that the self-employed underreport their income by about 30% of their 'true' income to the household surveys used in the analysis. They observe greater underreporting of income in the early part of the sample period and relate it to higher tax rates then. They also find evidence that the self-employed with higher education misreport their income to a lesser extent.

Mirus and Smith (1996) use the Canadian Family Expenditure Survey from 1990 and estimate that self-employed households conceal 12.5% of their 'true' total income. The

research on income tax non-compliance by self-employed households in Canada was continued by Schuetze (2002). He investigated a longer period, from 1969 to 1992, and estimated the share of underreporting for different years, demographic characteristics and occupations. The degree of non-compliance by self-employed households varies significantly with occupation, age and the number of household members that are self-employed. His estimations suggest that households which obtained 30% or more of their reported income from business activities concealed on average between 11% and 23% of the 'true' total household income.

There are two studies that cover the Nordic countries. Johansson (2005) estimates income underreporting by the self-employed in Finland for 1994–96. He finds that in households where only the head of the household was self-employed, on average 16.5% of the 'true' total income was not reported. In households in which at least two adults were self-employed, income was underreported by 42% on average. Engström and Holmlund (2009) hypothesise that the incentives for underreporting should be strong in a country with high tax rates on labour income and examine the connection between food expenditure and reported income in households in Sweden. They estimate that households with at least one self-employed member underreport their income by around 30%. They also distinguish between self-employed households with unincorporated and incorporated businesses, the latter of which must follow more stringent regulations. Engström and Holmlund (2009) conclude that underreporting is twice as prevalent among the self-employed who are unincorporated as among the self-employed with incorporated businesses.

Kim *et al.* (2009) obtain the permanent income component for estimating consumption propensities by eliminating transitory income fluctuations. They use panel data from the Korea Labour Income Panel Survey from 2000 to 2005 and the Russian Longitudinal Monitoring Survey from 1994 to 2000. Their approach leads to a point estimate of underreported income instead of the interval provided in other studies using the P&W method. They find that in Korea 38% and in Russia 47% of the 'true' total income of self-employed households is not reported. For reference, a list of studies using the P&W methodology is given in Table 1.

The P&W method has led to the development of alternative methods for estimating income underreporting. Lyssiotou *et al.* (2004) examine the use of non-parametric methods and propose a consumer demand system approach. They use the 1993 UK FES data and reach a larger share of underreporting than Pissarides and Weber (1989); the blue-collar self-employed leave 54% and the white-collar self-employed 39% of their 'true' income unreported.

Wangen (2005) develops two additional methods based on the P&W methodology to estimate income underreporting. His second method gives much wider intervals than the P&W method and indicates that actual business income is about 3.5 times higher than reported income for the UK. The large gap is in part the result of the model specification but it also suggests that refinements of the P&W methodology may lead to larger estimates of income underreporting by the self-employed.

Methodology

This study uses a modified version of the methodology pioneered by Pissarides and Weber (1989) to provide estimates of income not reported by households with business income.[2] The idea is to estimate a food Engel curve for all households but allow for a shift dummy for households with business income. After controlling for different household

Table 1. Studies using the methodology of Pissarides and Weber (1989) to estimate underreported income of self-employed households.

	Country	Database and time period	Definition of self-employment	Unreported income as share of 'true' total income
Pissarides and Weber (1989)	UK	Family Expenditure Survey 1982	Share of reported business income over 25%	White collar: 23–35%, blue collar 34–39%
Hurst *et al.* (2014)	USA	Consumer Expenditure Survey 1980–2003, Panel Study of Income Dynamics 1980–97	Reported self-employment	From CEX: 31%, from PSID: 29%
Schuetze (2002)	Canada	Family Expenditure Survey 1969–92 (of which, six years)	Share of reported business income over 30%	11–23%
Johansson (2005)	Finland	Household expenditure survey 1994–96	Reported self-employment	One person self-employed: 10–24%, two persons self-employed: 37–47%
Engström and Holmlund (2009)	Sweden	Household Budget Survey 1999–2004	Reported self-employment	With incorporated business: 15–20%, with unincorporated business: 40–50%
Kim *et al.* (2009)	Russia, South Korea	Longitudinal Monitoring Survey 1994–2000 (Russia), Labour Income Panel Survey 2000–05 (South Korea)	Reported self-employment	Russia: 47%, Korea: 38%

characteristics and wealth proxies that can induce differences in consumption behaviour, the propensity to food consumption is expected to be the same for both groups of taxpayers and the estimated shift dummy or gap will therefore reveal the share of earnings unreported by households with business income.

The crucial identifying assumption of the model is the attribution of the estimated gaps in food consumption to income underreporting. There are other potential explanations for an expenditure gap, such as heterogeneity in preferences, which may bias estimates of underreporting. Using food consumption with additional control variables to capture household heterogeneity is meant to address this problem. The risk of confusing heterogeneity with underreporting would be higher in a comparison of spending on durable goods, which have a wider variety of brands with larger differences in prices. Additionally, food typically cannot be classed as a business expense, which cars or telecommunication costs can, and when this happens, it is easier to report personal expenses as business expenses, possibly leading to the underreporting of consumption along with the underreporting of income.

According to the permanent income hypothesis, consumption is smoother than income as it is not affected by transitory income (Friedman 1957). This means that consumption depends on the permanent component of income and – in principle – not on current income. As the saving or dissaving of transitory income can be mistaken for misreporting, transitory income should be treated separately from permanent income.

The food Engel curve can be estimated using the following specification, where subscript i is the index for the households:

$$\log c_i = \alpha + \beta \log y_i^{\text{perm}} + X_i' \phi + \varepsilon_i. \tag{1}$$

The variable $\log c_i$ denotes the logarithm of food consumption, $\log y_i^{\text{perm}}$ is the logarithm of the permanent component of income, X_i is a column vector of control variables affecting consumption, and ε_i is an error term. The coefficient β is the income elasticity of food consumption and ϕ is a vector of the marginal effects of the control variables.

Usually only current income is reported in survey data, and this makes it difficult to distinguish between the permanent and transitory components of current income. The P&W methodology assumes that the current income deviates from the permanent income by a random variable that has a log-normal distribution across households. This assumption allows them to derive the permanent income component from current income. They also assume, realistically, that short-term fluctuations in current income are different for households of wage earners and households with business income as the latter group experiences more volatile income and this difference has to be taken into account in the estimations.

The Estonian HBS makes it possible to exclude the transitory component of income directly through the use of self-reported information.[3] We can therefore leave aside further assumptions regarding the distribution of the permanent and transitory components of the reported income.

It is assumed that household i might misreport its income by a factor κ_i. When income is misreported, the 'true' permanent component of income is thus

$$\log y_i^{\text{perm}} = \log \kappa_i + \log y_i^{\text{rep\_perm}}. \tag{2}$$

The variable $\log \kappa_i$ captures the unreported log permanent income that must be added to the log reported income, $\log y_i^{\text{rep\_perm}}$, to attain the 'true' log permanent income, $\log y_i^{\text{perm}}$. If $\kappa_i > 1$, the household underreports its 'true' permanent income.

The P&W model assumes that households of wage earners provide unbiased reporting of their income to the household budget survey, i.e. $\kappa_i = 1$ for the households of wage earners. Studies suggest that the share of wage income paid in the form of envelope wages is relatively small in Estonia although they contribute considerably to the livelihood of a small fraction of wage earners (Kriz *et al.* 2008, Meriküll and Staehr 2013). The assumption of P&W is not too restrictive in this respect: if the households of wage earners also systematically misreport their income to household surveys, the κ_i estimated for the households with business income would be an estimate of the relative difference in underreporting by households with business income and households of wage earners.

The random variable κ_i is assumed to have a log-normal distribution, and $\log \kappa_i$ can be expressed as the deviation from its mean:

$$\log \kappa_i = \mu_\kappa + \nu_i. \tag{3}$$

The term μ_κ is the mean of the logarithm of κ_i, and $\mu_\kappa > 0$ consequently entails underreporting by the average household with business income. The random variable ν_i has zero mean and constant variance $\sigma_\nu^2 > 0$ within each group. The variance for the group of households with business income is labelled $\sigma_{\nu|\text{B}}^2$, while the variance for the group of households of wage earners is labelled $\sigma_{\nu|\text{W}}^2$.

By combining Equations (1), (2) and (3), the following consumption function is obtained:

$$\log c_i = \alpha + \beta \log y_i^{\text{rep\_perm}} + \beta\mu_\kappa + X_i'\phi + \xi_i. \tag{4}$$

The term $\beta\mu_\kappa$ denotes the shift of the intercept of the Engel curve for households with business income compared to households of wage earners and $\xi_i = \varepsilon_i + \beta\nu_i$ is an error term. Since $E[\log y_i^{\text{rep\_perm}} \xi_i] \neq 0$, reported permanent income has to be instrumented. The first stage estimation takes the form

$$\log y_i^{\text{rep\_perm}} = Z_i'\delta + \omega_i, \tag{5}$$

where Z_i is a column vector of identifying instruments, δ is a vector of estimated coefficients and ω_i is the error term of the first stage income regression. The instrumented log reported permanent income is labelled $\log \hat{y}_i^{\text{rep\_perm}}$.

The 'true' permanent income of the average household with business income is found as the reported permanent income multiplied by κ. It follows from Equation (4) that the (average) underreporting factor κ can be found as

$$\kappa = \exp\left(\mu_\kappa + \frac{1}{2}\sigma^2_{\nu|\text{B}}\right). \tag{6}$$

The variance $\sigma^2_{\nu|\text{B}}$ is not known, so additional assumptions must be applied. In Equation (5) the residual ω_i contains unexplained variation in permanent income and the deviation of actual from reported income ν_i. Assuming that unexplained variation in permanent income has the same variance for both groups, the only difference in the variance of the error term ω_i in Equation (5) between the two groups stems from the variance in the log underreporting factor, and $\sigma^2_{\nu|\text{B}}$ can therefore be estimated as $\sigma^2_{\nu|\text{B}} = \sigma^2_{\omega|\text{B}} - \sigma^2_{\omega|\text{W}}$.

Unlike estimations undertaken using the P&W methodology in unaltered form, we obtain a point estimate of κ. There is no need to impose additional assumptions on the distribution of permanent income as the Estonian HBS makes possible the use of a self-reported measure of permanent income which omits transitory income fluctuations. This eliminates the need to report a range of estimates as the residual would include an additional error component (deviation of actual from permanent income) and the estimation of $\sigma^2_{\nu|\text{B}}$ would need an additional assumption regarding the covariance between the random variable of Equation (4) and the random variable of the permanent income process. Kim *et al.* (2009) also provide a point estimate but their method is based on statistical filtering of the current income variable.

Dataset and sample restrictions

This study uses monthly data from the Estonian Household Budget Survey (HBS), which is conducted by Statistics Estonia among a representative cross-section of Estonian households. We use the data from 2002 to 2007.[4] The dataset has previously been used by Kulikov *et al.* (2009) to investigate the saving behaviour of Estonian households and by Kukk *et al.* (2012) to estimate consumption sensitivities to shocks in income processes of different persistence. These studies contain detailed descriptions of the dataset.

Total consumption is the sum of 12 consumption categories. For the analysis we use the consumption of food, which includes eating outside the home. Wider consumption measures, such as non-durable and total consumption, include expenditure items like

telecommunication services, mobile telephones and computers that may be reported as business expenses and these measures are therefore not suitable for use in the estimation of income underreporting by households with business income.

The Estonian HBS makes it possible to extract two separate figures for monthly household income. The current after-tax household income contains five income categories: wage income, business-related income, property income, transfers and other income. In the questionnaire, business income is any type of earnings from self-employment and also earnings from activities that take place outside a regular wage contract, such as income from a start-up business or individual consultancy services. It includes income from the following sources: 1) registered self-employment activities, 2) provision of services, 3) self-production, 4) intermediation of products and services and 5) other business activities. It also contains dividends or any other kind of payment from a self-managed company – only dividends from investment with no active involvement in the company are regarded as capital income.

Additionally, the dataset includes the regular income, which is the household's assessment of its average monthly after-tax income when transitory income fluctuations are omitted. Kukk *et al.* (2012) show that the variable denoting the difference between current and regular income is indeed transitory income. The regular income cannot be considered to be fully permanent, but as pointed out by Pissarides and Weber (1989), it is important to exclude short-term fluctuations in income, while permanent income does not need to correspond one-to-one to the income in the permanent income hypothesis. The regular income is in all likelihood a good proxy for permanent income in the P&W model.

Different household characteristics that act as control variables and as instruments for income figures are nationality, gender, region of residence (Statistics Estonia's classification of five regions), the age of the household head, the number of children under 16 and a dummy variable indicating whether two adults are working. The dataset contains only a limited number of proxies of wealth, of which we use two: a dummy for renting housing rather than owning housing and a dummy for owning a second property.

We also use the variable temporary income shocks, which is defined as the difference between current and regular income reported in the Estonian HBS. Although theory hypothesises that consumption is insensitive to temporary income, Kukk *et al.* (2012) show that consumption does indeed react to temporary income shocks, although by substantially less than it does to regular income shocks. The variable will therefore be included in the consumption model to avoid omitted variable bias.

All income and consumption variables are in real terms. The nominal income and consumption series is deflated by the HICP index (with the index $= 1$ in 2005), while nominal food consumption is deflated by the HICP index for the food category.

Observations with missing income and food expenditure are excluded. Moreover, all observations where the economic activity status of the household head is classified as 'inactive' are also omitted from the sample. Finally, the sample is restricted to households with two adults, with or without children, since more precise estimates of the parameters can be obtained by focusing on a fairly homogeneous group. In the final sample there are 6016 cross-sectional observations.

The P&W methodology prescribes the division of the sample into a part deemed unlikely to underreport income and a part deemed likely to underreport. The sample can be divided in different ways but we make use of the method in Pissarides and Weber (1989), later used in Schuetze (2002). The method uses the share of reported business income in

Table 2. Split of the sample into different groups based on the share of reported business income in total income.

Share of reported business income in total income (%)	Number of households	Share of total households in the sample (%)
0	3085	51.3
0–5	1067	17.7
5–10	662	11.0
10–20	544	9.0
20–50	380	6.3
≥ 50	278	4.6
Total	6016	100.0

Source: Estonian HBS 2002–07.

reported total income as a means of identifying households most likely to underreport. A household is defined as self-employed if the share exceeds a given threshold. The approach makes it possible to investigate underreporting by households reporting different shares of business income.

Properties of the main variables

Table 2 shows the distribution of the share of reported business income to reported total income across households. Very large shares of households have both wage and business income and the balance between the two varies greatly between households. One adult may for instance be a wage-earner while the other one has business income or one person may have both wage and business income.

There are only a few households that report substantial business income. While almost half of the households in the sample (2931 households) report some positive business income, about 2/5 of this group report business income of less than 5%, while about 1/5 report business income of over 20% of their reported total income. In the case of underreporting of business income, this would make the share of reported business income in total income smaller than the 'true' share.

The P&W methodology provides estimates of underreporting of *total* income and the underreporting factor may therefore be different for households with different shares of business income. In our estimations we will focus on households whose reported business income is over 20% of reported total income but we also provide estimations for households where the share of business income is between 0 and 20%. We compare these households with business income with households of wage earners, defined as households with no business income.

The estimation of the Engel curve may lead to erroneous results if the share of business-related income is correlated with the regular income variable. If, for instance, households with higher regular income were to have a higher share of business income, the estimated results would be different for different income groups as the importance of the business income would vary across income groups. However, we do not find any evidence for this form of correlation. There is a weak negative correlation between regular income and the share of business income among households who receive business income (-0.18). If only households with more than 20% business income (the baseline sample) are considered, the correlation coefficient is 0.04. Figure 1 shows a scatter plot of the share of business-related income to log regular income and no systematic pattern is apparent.

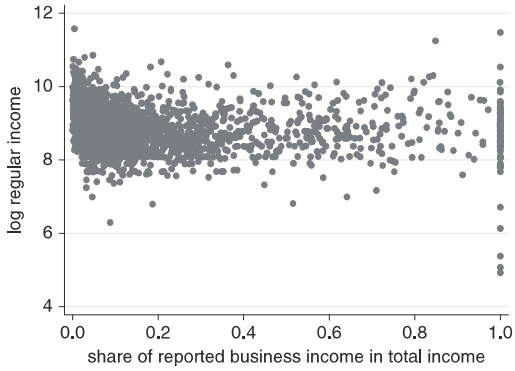


Figure 1. Scatter plot of share of business income and log regular income.

Table 3. Comparison of main variables for two household groups.

	Wage earners (no reported business income)		Reported business income ≥ 20% of reported total income	
Variable	Mean	St. dev.	Mean	St. dev.
Log food consumption[a]	7.638	0.533	8.044	0.537
Log regular income[a]	8.991	0.548	8.747	0.653
No. of obs.	3085		658	

Source: Estonian HBS 2002–2007.
Note: [a] The variables are expressed in 2005 prices. During the sample period, the *kroon* (EEK) was the currency in Estonia; the exchange rate was fixed at 15.65 EEK = 1 EUR.

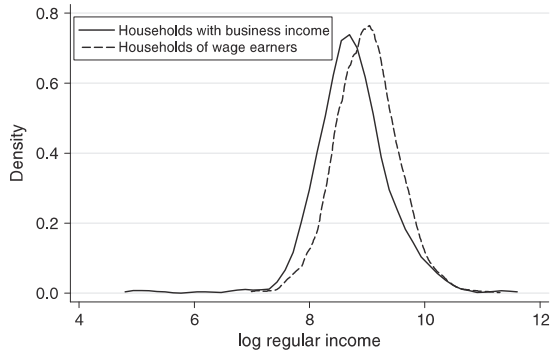


Figure 2. Kernel density functions of reported total income.
Note: Households of wage earners are households that do not report any business income; households with business income are households that report business income above 20% of reported total income. Kernel = epanechnikov, bandwidth = 0.1315.

Table 3 gives the main descriptive statistics of the income and consumption measures for households of wage earners (with no business income) and households with business income above 20% of reported total income. Households of wage earners have a lower mean food consumption than households with business income, but report a higher level of regular income. Given the same propensity to food consumption for the two groups and that food consumption is correctly reported, the gap can be explained by income underreporting by the households with business income.

The kernel density functions in Figure 2 show that the distributions of food consumption and regular income for wage earners and for households with business income are very similar while the means of the distributions are different. As the distributions are so similar, it is reasonable to use the P&W method to estimate the gap between the two groups.

The empirical model and estimation results

We estimate the following Engel curve, which is the empirical equivalent of Equation (4):

$$\log c_i = \alpha + \beta \log \hat{y}_i^{\text{rep\_reg}} + \gamma D_i + X_i'\phi + \xi_i. \tag{7}$$

The variable $\log c_i$ is the log of household food consumption, $\log \hat{y}_i^{\text{rep\_reg}}$ is the instrumented log reported regular income and D_i is a dummy variable which in the baseline estimation takes the value one for households with business-related income of over 20% of reported total income. The control variables consist of a number of household characteristics captured in a column vector X_i which includes time fixed effects for the 71 months in the sample. Household characteristics include the log temporary income, age and age squared of the household head, the number of children under 16, a dummy for there being two income earners, a dummy for ownership of a second property and a dummy for renting the main residence. The dataset lacks wealth variables that may affect consumption along with the income variable (Attanasio 1999). Nevertheless, inclusion of dummies for real estate should capture the wealth effect on consumption as real estate represents the main share of the wealth of households (ECB 2013).[5]

We use the household head's education level, gender and nationality in addition to regional dummies as additional instruments for regular income. According to the standard Mincer model, the education level is an important determinant of income level (Mincer 1976). Significant income gaps are also observed between genders and nationalities in Estonia (Leping and Toomet 2008, Anspal and Rõõm 2011). It is assumed that the estimated coefficients of the control variables and of the instruments for the income variable $\log \hat{y}_i^{\text{rep\_perm}}$ do not differ between households of wage earners and households with business income.

The dummy variable D_i in Equation (7) replaces $\beta\mu_k$ in Equation (5), i.e. the estimated coefficient γ is equivalent to $\beta\mu_\kappa$ in the theory model.[6] The underreporting factor κ indicates how much the regular income reported by households with business income must be scaled up in order for these households to reach the same propensity to food consumption as households of wage earners:

$$\kappa = \exp\left(\frac{\gamma}{\beta} + \frac{1}{2}\left(\sigma^2_{\omega|\text{B}} - \sigma^2_{\omega|\text{W}}\right)\right). \tag{8}$$

The underreporting factor κ denotes the 'true' total income in relation to the reported total income of the households with business income. The term $(\kappa - 1)/\kappa$ is correspondingly the underreported income as a share of the 'true' total income. The underreporting factor κ

is computed using estimates of $\sigma^2_{\omega|B}$ and $\sigma^2_{\omega|W}$. A simplified measure κ_S can be computed if these terms are omitted (cf. Hurst *et al.* 2014):

$$\kappa_S = \exp(\gamma/\beta). \tag{9}$$

The simplified underreporting factor κ_S is typically a good approximation of the standard factor κ. The simplified factor κ_S facilitates statistical hypothesis testing since its standard error can easily be calculated.

The results of the estimation of the food Engel curve in Equation (6) are given in Table 4. Here we present the variables of main interest, i.e. the coefficients of the regular income variables and the business income dummy. The estimations for the full baseline model and for the first stage regression are given in Table A.1 in the Appendix.

Column (1) in Table 4 shows the results of the baseline model where a dummy for households with reported business income over 20% is included. The households for which the share of reported business income is between 0 and 20% are excluded from the sample altogether, hence households with no business income behave as a comparison

Table 4. Food consumption estimations.

	(1) Baseline (reported business income ≥ 20%)	(2) Two groups of households with business income	(3) Five groups of households with business income
β (consumption propensity)	0.612***	0.611***	0.625***
	(0.039)	(0.039)	(0.031)
γ (business income ≥ 20%)	0.546***	–	–
	(0.025)		
γ_1 (business income ≥ 50%)	–	0.545***	0.546***
		(0.028)	(0.028)
γ_2 (business income 20–50%)	–	0.548***	0.581***
		(0.039)	(0.039)
γ_3 (business income 10–20%)	–	–	0.444***
			(0.019)
γ_4 (business income 5–10%)	–	–	0.290***
			(0.017)
γ_5 (business income < 5%)	–	–	0.131***
			(0.016)
R^2	0.297	0.298	0.331
Endogeneity test	36.23	36.66	50.94
[*p*-value]	[0.000]	[0.000]	[0.000]
Hansen *J*-test	13.44	13.45	18.69
[*p*-value]	[0.062]	[0.062]	[0.001]
No of obs.	3743	3743	6016
R^2 of the first regression	0.453	0.454	0.449
F-stat of the first regression	40.49	40.16	57.54
[*p*-value]	[0.000]	[0.000]	[0.000]

Notes: IV estimations with GMM estimator. Education level, nationality, gender and five regional dummies are used as instruments. Log temporary income, age and age squared of the household head, the number of children, two income earners, ownership of a second property, renting the main residence and 71 monthly time fixed effects are included in the estimations, but the results are not shown in the table. Robust standard errors are reported in round parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively. *F*-stat is the test statistic for the Wald test of overall goodness of the fit of the first regression.

group. The estimated coefficients are statistically significant at the 1% level. The propensity to consumption of food, β, is 0.61, while the coefficient of the business income dummy, γ, is 0.55. For the first stage regression the F-statistic for the overall goodness of fit is 40.49 and the coefficient of determination R^2 is 0.45 and this suggests that the chosen variables are appropriate instruments of the regular income variable.

As a robustness check we estimate the model with different sets of control variables. The results are shown in Table A.2 in the Appendix. The estimation results are very stable regardless of the number and choice of control variables. The results confirm that the propensity to consume food is markedly higher for the households with business income than for the households of wage earners.

The estimated underreporting factor refers to underreporting of the *total* income of the household, not only of business income. Different shares of business income in reported total income may therefore lead to different estimates of the underreporting factor. This issue is examined in more detail in the estimations in columns (2) and (3) in Table 4. Column (2) shows the coefficients of two different dummies, one for which the reported business income is in the interval 20–50% and one for which the reported business income is over 50%, but the estimated coefficients, γ_1 and γ_2, are very similar. Column (3) shows the results when dummies for households with lower shares of reported business income are included. The lower the share of business income is, the lower the estimated coefficients of the dummy variables γ_3, γ_4 and γ_5 are, which is unsurprising given that the estimated lower dummy coefficient indicates lower underreporting of *total* income. The results in columns (2) and (3) made us choose the 20% cut-off of business income in reported total income for our baseline scenario.

Table 5 shows the underreporting factors κ and κ_S computed from the estimated coefficients in Table 4 and the residual variances from the first stage of the IV regression, cf. Equations (8) and (9). Column (1) shows that the simple underreporting factor, κ_S, is around 2.4 and the factor is estimated very precisely. The standard underreporting factor taking into account the first stage variances, κ, is 2.6. Focusing on the latter, the reported income for households with reported business income over 20% must be multiplied by 2.6 to reach the same propensity to food consumption as households of wage earners. Put differently, households with reported business income over 20% have left unreported $1.631/2.631 \approx 62\%$ of their 'true' total income.

It is important to emphasis that the 'true' income in this context is relative to the income reported by wage earners. As discussed above, it is conceivable that households of wage earners receive envelope wages and misreport their income to the HBS. We can only calculate the gap in underreporting between households with business income and households of wage earners, so the computed underreporting of the 'true' income should be seen as relative to the underreporting by households of wage earners.

Table 5. Estimates of income underreporting by households with business income.

	(1) Reported business income $\geq$ 20%	(2) Reported business income 10–20%	(3) Reported business income 5–10%
Underreporting factor κ_S (with standard errors)	2.442	2.035	1.591
	(0.137)	(0.080)	(0.053)
Underreporting factor κ	2.631	2.005	1.561
Share of 'true' total income unreported	62.0%	50.1%	35.9%

Notes: Authors' calculations based on the results in Table 4. Column (1) is estimated using the coefficients of column (1) in Table 4. Columns (2) and (3) are estimated using the coefficients of column (3) in Table 4.

The underreporting factors in columns (2) and (3) in Table 5 are computed using the regression results from column (3) in Table 4, including the consumption propensity β and, respectively, the business income dummies γ_3 and γ_4. For the households that have reported business income of less than 20%, the income variance does not differ from the income variance of wage earners and hence the two underreporting factors κ_S and κ are very similar. In order to reach the same propensity to food consumption as households of wage earners, the income must be multiplied by 2.0 for households where reported business income is 10–20% or by 1.6 for households where it is 5–10%.

These results indicate considerable underreporting of *total* income in households in which the share of business income is relatively small. If it were assumed that business income is the main source of the underreporting, the estimations would imply very substantial underreporting of business income. For households with reported business income of less than 20% we would need to multiply their *business income* by a factor of at least six to obtain the given underreporting factor in *total* income. We cannot, however, assume that the source of the underreporting is only business income; the results suggest that households who have business income are more prone to hiding all sources of income too. The estimations imply that even small shares of reported business income are a good indicator of *total* income underreporting.

The results in Table 5 suggest that the extent of underreporting by households with business income is very substantial in Estonia. The baseline model implies that 62% of the income is left unreported, but the share of unreported income would be somewhat smaller if a lower cut-off level than 20% were chosen.

The corresponding results from other countries discussed earlier generally show lower shares of underreporting. Among developed countries the upper limit of the share of underreporting is estimated to be around 30–40%. Estonia is a transition country with particular economic and institutional structures, and the substantially higher estimates for Estonia than for developed countries are consistent with the estimates of the shadow economy in Estonia and other European transition economies. This applies to the overall size of the shadow economy, where the estimates for the transition countries are typically double or triple the estimates for developed economies (Tafenau *et al.* 2010, Schneider *et al.* 2010). It also applies to the payment of undeclared wages and the extent of unreported employment (Williams 2008, 2009).

Among the group of transition countries, the only study using the P&W method is for Russia, where the result is underreporting of 47% of 'true' total income, i.e. a slightly smaller share than was found in this study (Kim *et al.* 2009). The study on Russian data, however, uses a different definition of self-employed households, so the results are not directly comparable. Moreover, the computed underreporting factor for Russia has wide confidence intervals.

The results for Estonia are not *fully* comparable with those of other studies that use the P&W methodology. Previous studies have typically defined some households as wage earners even if they report earning some business income. We find that households with business income underreport their income even if the share of reported business income in reported total income is small (Table 5, columns (2) and (3)).

From the findings in Table 5, we define a household of wage earners as a household that reports *no* business income; the comparison group in previous studies is thus more mixed than it is in our estimations. When we use a similar definition to those in previous studies, we obtain somewhat lower but still qualitatively similar results. If we use the same cut-off level of 25% reported business income as used in Pissarides and Weber (1989) for instance, and include households with reported business income of lower than 25% in the

sample of wage-earners, the estimated underreporting parameter for Estonia would be 2.37, which entails a share of unreported income equal to 57.8%. Nevertheless, we prefer to define households of wage earners as those who report no business income precisely because we find evidence of underreporting among households with a share of reported business income between 0 and 20%.

The underreporting by households with business income found in this study can in fact originate from two sources. One way for households to underreport is to leave some business income unreported, while the other is to over-report expenses related to business activities. This is because the tax code allows households to deduct their business-related expenses before reporting the business income to the tax authorities. This lets households report expenses for car maintenance, telecommunication and even housing (if the office is at home) as business expenses, thus lowering the reported business income. This type of tax behaviour is arguably used quite extensively in Estonia.

The gap between the consumption propensities for households with business income and households of wage earners could be due to factors other than underreporting of income. If households involved in business activities have preferences that lead to a different consumption pattern than that of households of wage earners, then the estimations could also capture this effect. As argued above, this is considerably less likely for food consumption than for non-durables or durables, for which consumption can vary more across different goods and brands. Moreover, we employ a large number of control variables to take account of heterogeneity in tastes across households.

We also investigate the dynamics of the underreporting factor over time by estimating the model separately for 2002–03, 2004–05 and 2006–07. This exercise is particularly pertinent as the period was characterised by rapid economic and institutional change (Meriküll and Staehr 2010). Estonia joined the EU in 2004 and many reforms were passed in the years before then to achieve compliance with the *acquis communautaire*. The result was numerous changes in company law, taxation rules, industrial policy and government institutions. The country saw rapid development with annual rates of economic growth above 6%. The flat tax rate on personal and corporate income was 26% until 2004 but was then gradually lowered to 22% in 2007.

The estimations are for households that have business income over 20%. The results in Table 6 provide some evidence for parameter changes across time as the underreporting factor is lower for the period after 2004, but the difference is not statistically significant. In any case, the share of 'true' income not reported is very large in all three subsamples; it is relatively constant across the subsamples and with no clear direction of change. This suggests that EU accession, institutional development and rapid economic change did not materially affect the extent of underreporting by households with business income.

Table 6. Estimates of income underreporting in different periods.

	(1) 2002–03	(2) 2004–05	(3) 2006–07
Underreporting factor κ_S (with standard errors)	2.755 (0.275)	2.241 (0.186)	2.483 (0.327)
Underreporting factor κ	3.130	2.337	2.569
Share of 'true' total income unreported	68.1%	57.2%	61.1%
No. of observation.	1411	1063	1269

Notes: Authors' calculations.

We also produced estimations for different subsamples but again the wide confidence intervals did not allow us to draw any firm conclusions. We did not find any statistically significant difference between the coefficient estimates for households of different educational levels or for households living in different regions (not reported) but the results hinge on a small number of observations in the different subsamples.

Final comments

This article estimates the extent of income underreporting to the Estonian Household Budget Survey in 2002–07 by households with business-related income relative to households of wage earners. The analysis uses the methodology pioneered by Pissarides and Weber (1989) but modifies it to take into account the availability of a self-reported measure of regular or permanent income in the Estonian HBS. If individuals provide consistent data about their income to all data collectors, the underreporting results based on data from the HBS may also be taken as rough proxies of income underreporting to tax authorities (Hurst *et al.* 2014).

The baseline estimation considers income underreporting by households for which reported business income comprises 20% or more of total income. The result is that the reported total income must be multiplied by 2.6 to reach the same propensity to food consumption as households of wage earners. In other words, households with business income over 20% have left unreported 62% of their 'true' income, i.e. the sum of reported and unreported total income. The 'true' income in this context is relative to the income reported by households of wage earners.

The underreporting of income is somewhat lower when the share of reported business income is less than 20% but it is still substantial; the income should be multiplied by 2.0 for households where business income is 10–20% and by 1.6 for households where it is 5–10%. The sample period 2002–07 was characterised by rapid economic development, EU membership in 2004 and rapid institutional changes, but no trend in the underreporting is apparent when the sample is split into three time subsamples.

The underreporting results for households with business income are somewhat higher for Estonia than those in most other studies using the methodology of Pissarides and Weber (1989). This applies particularly when compared with developed economies, but less so when Estonia is compared with Russia, another transition economy. Studies of the extent of envelope wages and the overall size of the shadow economy show a similar pattern, i.e. that the extent of unreported activity is much larger for transition economies than for developed economies.

The estimates may represent an upper bound for the underreporting of income, given that some fraction of the difference in food consumption may originate from sources other than underreporting. The upshot is, nevertheless, that underreporting by households with business income is very pronounced. If households are consistent in their reporting to the household survey and to the tax authorities, the estimated income underreporting is a possible indicator of the level of tax evasion and tax avoidance by Estonian households with business income. The present analysis suggests that detailed studies using tax and register data would indeed be a worthwhile undertaking.

Acknowledgements

The authors would like to thank Arvo Valtin from Statistics Estonia for providing the underlying data and Dmitry Kulikov for arranging the database. They also thank Jens Hölscher, Jaanika Meriküll, Alari Paulus, Tairi Rõõm and participants at the ECEE4 conference in Tallinn and the 2013 EEA

conference in Gothenburg for valuable comments on earlier versions of this article. This work was supported by [grant number SF0140059s12]. The views expressed are those of the authors and not necessarily those of Eesti Pank.

Notes

1. Estimates of income underreporting by households with business income are particularly important because many other countries seek to encourage entrepreneurship and self-employment. Estonia provides support for the unemployed to assist them in establishing businesses and becoming self-employed (Leetmaa and Nurmela 2010).
2. The consumer demand system approach by Lyssiotou *et al.* (2004) is an alternative approach, but the P&W methodology is the most widely used and the use of the P&W methodology therefore facilitates comparison with the results for other countries.
3. The survey includes questions on current income and regular income. The difference can be considered as transitory income. Further investigation of the two income components is provided in Kukk *et al.* (2012).
4. The data collection was discontinued in 2008 due to budgetary constraints at Statistics Estonia.
5. Additionally, instrumentation should reduce or eliminate the problem of biased estimates of the income coefficient due to the omission of relevant wealth variables (Wooldridge 2002).
6. If the business-related income is taken as endogenous, the dummy variable D_i would need to be instrumented. Experimentation with different instruments from the dataset did not produce satisfactory results, perhaps due to weak instruments, and we therefore decided not to instrument D_i.

References

Anspal, S. and Rõõm, T., 2011. *Gender pay gap in Estonia: empirical analysis*. Research Report for the Ministry of Social Affairs of Estonia [online]. Available from: http://www.sm.ee/fileadmin/meedia/Dokumendid/V2ljaanded/Publikatsioonid/2011/Gender_pay_gap_Estonia_analysis.pdf

Attanasio, O., 1999. Consumption. *In*: J. Taylor and M. Woodford, eds. *Handbook of macroeconomics*. Rotterdam: Elsevier, 741–812.

Black, T. *et al.*, 2012. Federal tax compliance research: tax year 2006 tax gap estimation US Department of the Treasury. *Internal Revenue Service research, analysis & statistics working paper,* March.

Bruce, D., 2000. Effects of the United States tax system on transitions into self-employment. *Labour economics*, 7 (5), 545–574.

ECB, 2013. The Eurosystem Household Finance and Consumption Survey. Results from the first wave. *European Central Bank, Statistics paper series,* no. 2.

Engström, P. and Holmlund, B., 2009. Tax evasion and self-employment in a high-tax country: evidence from Sweden. *Applied economics*, 41 (19), 2419–2430.

Eurobarometer, 2007. Undeclared work in the European Union. *Special Eurobarometer,* no. 284 [online], European Commission. Available from: http://ec.europa.eu/public_opinion/archives/ebs/ebs_284_en.pdf

Friedman, M., 1957. *A theory of the consumption function*. Princeton and Oxford: Princeton University Press.

Hurst, E., Li, G. and Pugsley, B., 2014. Are household surveys like tax forms: evidence from income underreporting of the self employed. *Review of economics and statistics*. 96 (1), 19–33.

Johansson, E., 2005. An estimate of self employment income underreporting in Finland. *Nordic journal of political economy*, 31 (1), 99–109.

Kim, B., Gibson, J. and Chung, C., 2009. Using panel data to exactly estimate income underreporting of the self employed. *Labor economics working papers,* no. 22996, East Asian Bureau of Economic Research.

Kriz, K. *et al.*, 2008. Why do individuals evade payroll and income taxation in Estonia? *In*: M. Pickhardt and E. Shinnick, eds. *Shadow economy, corruption and governance*. Cheltenham: Edward Elgar Publishing.

Kukk, M., Kulikov, D. and Staehr, K., 2012. Consumption sensitivities in Estonia: income shocks of different persistence. *Bank of Estonia working papers,* no. 3/2012.

Kulikov, D., Paabut, A. and Staehr, K., 2009. A microeconometric analysis of household saving in Estonia: income, wealth and financial exposure. *In*: D. Mayes, ed. *Microfoundations of economic success: lessons from Estonia*. Cheltenham: Edward Elgar, 190–241.

Leetmaa, R. and Nurmela, K., 2010. EEO review: self-employment, 2010, Estonia [online]. European Employment Observatory. Available from: http://www.eu-employment-observatory.net/resources/reviews/Estonia-SERvw2010.pdf

Leping, K.O. and Toomet, O., 2008. Emerging ethnic wage gap: Estonia during political and economic transition. *Journal of comparative economics*, 36 (4), 599–619.

Lyssiotou, P., Pashardes, P. and Stengos, T., 2004. Estimates of the black economy based on consumer demand approaches. *Economic journal*, 114 (497), 622–640.

Meriküll, J. and Staehr, K., 2013. Unregistered production and employment in Estonia: measurements and developments. *Journal of applied economics and business*, 1 (2), 23–34.

Meriküll, J. and Staehr, K., 2010. Unreported employment and envelope wages in mid-transition: comparing developments and causes in the Baltic countries. *Comparative economic studies*, 52 (4), 637–670.

Mincer, J., 1976. *Schooling, experience, and earnings*. New York: NBER Press.

Mirus, R. and Smith, R., 1996. Self-employment, tax evasion, and the underground economy: micro based estimates for Canada. *Working paper,* no. 1002, International tax program, Harvard Law School.

Pissarides, C. and Weber, G., 1989. An expenditure based estimate of Britain's black economy. *Journal of public economics*, 39 (1), 17–32.

Schneider, F., 2011. The shadow economy and shadow economy labor force: what do we (not) know? *Discussion paper,* no. 5769, Institute for the Study of Labor (IZA).

Schneider, F., Buehn, A. and Montenegro, C.E., 2010. New estimates for the shadow economies all over the world. *International economic journal*, 24 (4), 443–461.

Schuetze, H., 2002. Profiles of tax noncompliance among the self employed in Canada: 1969–1992. *Canadian public policy*, 28 (2), 219–237.

Slemrod, J., 2007. Cheating ourselves: the economics of tax evasion. *Journal of economic perspectives*, 21 (1), 25–48.

Torgler, B., 2003. Tax morale in transition countries. *Post-communist economies*, 15 (3), 357–381.

Tafenau, E., Herwartz, H. and Schneider, F., 2010. Regional estimates of the shadow economy in Europe. *International economic journal*, 24 (4), 629–636.

Wangen, K.R., 2005. An expenditure based estimate of Britain's black economy revisited. *Discussion papers,* no. 414, Research Department of Statistics Norway.

Williams, C.C., 2008. Envelope wages in Central and Eastern Europe and the EU. *Post-communist economies*, 20 (3), 363–376.

Williams, C.C., 2009. The commonality of envelope wages in Eastern European economies. *Eastern European economics*, 47 (2), 37–52.

Wooldridge, J.M., 2002. *Econometric analysis of cross section and panel data*. Cambridge MA: MIT Press.

Appendix

Table A.1. Full estimations of IV regression model.

Dependent variable:	First stage IV regression $\log y_i^{\text{regular}}$	Second stage IV regression $\log c_i$
Log regular income (β)		0. 612***
		(0.039)
Household education level 2	0.116***	–
	(0.032)	
Household education level 3	0.325***	–
	(0.033)	
Non-Estonian	−0.226***	–
	(0.019)	
Female	−0.160***	–
	0.015	
Region Northeast	−0.251***	–
	(0.021)	
Region Central	−0.183***	–
	(0.026)	
Region West	−0.219***	–
	(0.026)	
Region South	−0.208***	–
	(0.022)	
Business income dummy (γ)	−0.212***	0.546***
	(0.025)	(0.025)
Log temporary income	−0.076***	0.170***
	(0.015)	(0.016)
Age − mean age	−0.001**	0.002*
	(0.001)	(0.001)
(Age − mean age)^2	0.0002***	−0.0002***
	(0.000)	(0.000)
Number of children	0.086***	0.087***
	(0.010)	(0.010)
Two income earners	0.333***	−0.057***
	(0.021)	(0.022)
Owning second property	0.107***	−0.050**
	(0.025)	(0.025)
Renting main residence	−0.077***	−0.051*
	(0.026)	(0.029)
Constant	8.447***	2.262***
	(0.082)	(0.331)
R^2	0.453	0.297
F-stat	40.49	–
[p-value]	[0.000]	
No. of obs.	3743	3743

Notes: IV estimations. 71 monthly dummies are included in the estimations but are not shown in the table. Robust standard errors are reported in round parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively. F-stat is the test statistic for the Wald test of overall goodness of the fit of the first regression.

Table A.2. Robustness test of the regression to different sets of control variables. Dependent variable: $\log c_i$.

	(1)	(2)	(3)	(4)	(5)
Log regular income (β)	0.602***	0.618***	0.603***	0.604***	0.604***
	(0.032)	(0.037)	(0.038)	(0.038)	(0.038)
Dummy for business income (γ)	0.555***	0.583***	0.551***	0.547***	0.547***
	(0.026)	(0.025)	(0.025)	(0.025)	(0.025)
Temporary income and two income earners	No	Yes	Yes	Yes	Yes
Age and number of children	No	No	Yes	Yes	Yes
Renting main residence and owning second property	No	No	No	Yes	Yes
Time fixed effects	No	No	No	No	Yes
R^2	0.213	0.244	0.278	0.278	0.278
No. of obs.	3743	3743	3743	3743	3743

Notes: IV estimations. Education level, nationality, gender and regional dummies are used as instruments. Robust standard errors are reported in parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% level respectively.

Appendix 4. Publication IV

Reprint from:

Kukk, Merike, and Karsten Staehr (forthcoming). Identification of households prone to income underreporting: employment status or reported business income? *Public Finance Review*, available online from 1 December 2015 at: http://pfr.sagepub.com/, DOI: 10.1177/1091142115616182

Article

Public Finance Review
1-29

DOI: 10.1177/1091142115616182
pfr.sagepub.com
SAGE

Identification of Households Prone to Income Underreporting: Employment Status or Reported Business Income?

Merike Kukk[1] and Karsten Staehr[1,2]

Abstract

Pissarides and Weber propose using data on income and food consumption for estimating the extent of income underreporting by the self-employed, a group seen to be prone to income underreporting. This paper is the first to investigate the importance of the way in which these households are identified in such analyses. Using household budget data from Estonia, different ways are used to identify households prone to income underreporting and to estimate the extent of the underreporting. The share of unreported income is estimated to be substantially larger when underreporting households are identified using their share of reported business income than when they are identified using their employment status.

[1] Department of Finance and Economics, Tallinn University of Technology, Tallinn, Estonia
[2] Economics and Research Department, Eesti Pank, Tallinn, Estonia

Corresponding Author:
Merike Kukk, Department of Finance and Economics, Tallinn University of Technology, Akadeemia tee 3-478, Tallinn 12618, Estonia.
Email: merike.kukk@ttu.ee

Further analysis combines the different identification methods and reveals that the employment status provides no information on underreporting when the share of business income is taken into account. The share of reported business income is the most informative indicator of underreporting.

Keywords
income underreporting, business income, self-employment, engel curve

Pissarides and Weber (1989) proposed a new methodology for estimating the extent of income underreporting by the self-employed, a group of taxpayers that are usually seen as prone to underreporting their income. The main idea is to compare the relation between reported consumption and income for the self-employed and for people that are not self-employed. This paper is the first to examine the importance of the way the self-employed, i.e. the households prone to underreporting, are identified in empirical studies using the methodology of Pissarides and Weber (1989).

It is typically easier for the self-employed to underreport their income to tax and statistics authorities than it is for wage earners. Information on the income of employed individuals is available to the authorities through third-party reporting, while the self-employed typically have substantial discretion over the income reported as audits of income and deductions are infrequent or imprecise (Soos 1990). These differences suggest that the self-employed may underreport income to a larger extent than households of wage earners. Bruce (2000) posits that the chance to underreport represents an incentive for people to choose to be self-employed. The experimental study of Alm, Deskins, and McKee (2009) shows that individuals with a higher share of income not reported to the tax authority by a third party exhibit lower tax compliance.

It is challenging to estimate the extent of underreporting by the self-employed. Pissarides and Weber (1989) estimate the extent of underreporting by using information on food consumption and income from the UK household budget survey. The underlying idea is that households are likely to report their consumption expenses relatively precisely as the households produce detailed consumption accounts for the survey, while some households may underreport their income to the statistics office. The extent of income underreporting by the self-employed is then estimated from an Engel food consumption function that includes a shift term for

self-employed households. The main identifying assumption is that self-employed households and other types of households, ceteris paribus, have the same propensity to consume out of their total (reported and unreported) income.

Underreporting of income to a statistics office does not necessarily entail underreporting to the tax authority. Pissarides and Weber (1989) argue, however, that households may seek not to report higher income to a household budget survey than they do to the tax authority as the survey respondents may fear that information is exchanged between the statistics office and the tax authority. Paulus (2015a) finds, using data for 2007 and 2008, a substantial overlap between the reporting of households to the Estonian Social Survey and to the Estonian tax authorities. The results may thus be seen as an indication of possible tax evasion by the self-employed.

Pissarides and Weber (1989) use data from the 1982 Family Expenditure Survey for the UK and find that the share of total (reported and unreported) income not reported by the self-employed is 23 to 39 percent. The estimated income underreporting is *relative to the comparison group*, meaning that the estimate depicts the underreporting of self-employed households *on top of* the possible underreporting by the comparison group. In many cases, however, households that only have income from employment would not easily be able to underreport their income because of third-party reporting (Pissarides and Weber 1989).

After a hiatus, a number of studies have been published using the Pissarides and Weber (1989) methodology for different countries. Schuetze (2002) uses Canadian data for six years within the time period 1969 to 1992 and finds that self-employed households left 11 to 23 percent of income unreported relative to the comparison group of wage earners. Johansson (2005) uses data from the 1994 to 1996 Finnish household budget surveys and estimates that the share of unreported income is 10 to 24 percent if one person in the household is self-employed and 37 to 47 percent if two people are self-employed. Hurst, Li, and Pugsley (2014) use data from two different surveys from the USA covering around 20 years and reach very similar results from both data sources, finding that households that state their employment status as self-employed underreport around 30 percent of their income.

Gibson, Kim, and Chung (2009) are the first to use data from countries outside Western Europe and North America. They estimate the underreporting share of the self-employed in South Korea to be around 38 percent (data for 2000–2005), while the underreporting share in Russia amounts to

47 percent (data for 1994–2000). Kukk and Staehr (2014) use data from Estonia for 2002 to 2007 and find that the underreporting share of the self-employed in Estonia is around 62 percent of their total reported and unreported income.

Two other papers are of particular relevance for this paper. Engström and Holmlund (2009) consider underreporting in Sweden using data for 1999 to 2004. They find an underreporting share of 15 to 20 percent for self-employed households with an incorporated business and 40 to 50 percent for households with an unincorporated business. Paulus (2015b) uses data from the Estonian Social Survey and Estonian tax records for 2007 and 2008 and finds an underreporting share by the self-employed of 56 percent when the comparison group consists of *wage earners in the public sector*. These papers suggest that different classifications of the self-employed and the comparison group may be important for the results.

The Pissarides and Weber (1989) methodology prescribes the division of the sample into two groups: those deemed prone to underreporting income and the comparison group deemed less likely to underreport income. Two different methods for dividing the sample are used in the literature. One method considers the *share of reported business income* in reported income and defines a household as prone to underreporting if the share exceeds a given threshold. The other method uses the *employment status* reported by the household and defines a household as prone to underreporting if self-employment is the reported employment status. The first method, based on the share of reported business income, is used in Pissarides and Weber (1989), Schuetze (2002), and Kukk and Staehr (2014). The second method, based on the reported employment status, is used in Hurst, Li, and Pugsley (2014), Johansson (2005), Engström and Holmlund (2009), and Paulus (2015b).

Pissarides and Weber (1989) are the only ones to discuss the identification choice. They state that they reach similar results when using either of the two methods to identify households prone to underreporting, but they only provide the results when the households are identified using the reported share of business income. All other studies use one of the identification methods and do not investigate whether the particular choice is important for the results, probably because of data availability. Although household budget surveys may contain data on both reported employment status and business income, income data are typically not collected at a detailed level and data on business income may therefore not be available.[1] Moreover, in some household budget surveys the time periods of the consumption and income data do not match.[2]

This paper is the first to present a detailed analysis of the consequences of the choice of method for identifying the households that may be prone to income underreporting. The Estonian Household Budget Survey (HBS) for the period 2002 to 2007 is uniquely suited for this analysis because the survey contains questions that make it possible to use both identification methods. The survey collects detailed information both on consumption and on different income sources and there is no mismatch between the periods of consumption and income data.

The Tax System and Tax Compliance in Estonia

Estonia regained its independence from the Soviet Union in 1991. It embarked on an ambitious reform process and by the end of the 1990s it exhibited a market economy with a small public sector and a low tax intake relative to GDP (Staehr 2004). Fiscal policy is prudent and in rankings of fiscal management Estonia is typically placed alongside or above most Western European countries (Fabrizio and Mody 2010). The policy of e-governance has let individuals and firms file their taxes online since 2000.

The Estonian tax system is simple (Staehr 2009). Personal income from employment or business activities is subject to two main taxes, the social tax or payroll tax and the personal income tax. The social tax, amounting to 33 percent of personal income, is levied without any exemptions. The personal income tax is at a flat rate levied on taxable income, (i.e. personal income minus the basic exemption and a few other deductions including interest payments on housing loans, educational expenses, charitable donations and some pension contributions). Taxable income is subject to a flat tax rate initially set at 26 percent but reduced gradually from 2005.

The main difference between the tax treatment of personal income from employment and personal income from business activities is that the former is subject to third-party reporting, while this is not the case for income from business activities. Revenue and deductible business expenses are reported to the tax authorities by the taxpayer and typically without any third-party reporting. This makes it fairly easy to underreport personal income by leaving out revenue or by overstating business deductions. Overstating of deductions is aided by the fact that in some cases there is no clear distinction between business expenses and personal expenses. It may for instance be possible to report housing costs, office equipment and car purchases as business expenses.

Different studies provide quite a broad range of estimates of the size of the shadow economy in Estonia, an outcome that partly reflects the

difficulty of providing estimates of activities that are kept in the shadow. The estimates are generally comparable to those for other EU countries from central and eastern Europe and southern Europe (Schneider and Williams 2013; Tafenau, Herwartz, and Schneider 2010).

Data

The empirical analyses in this paper are based on data from the 2002 to 2007 versions of the Estonian Household Budget Survey. The end point of the sample is dictated by the fact that the data collection was halted in 2008–2009 and when it resumed in 2010 it started following the guidelines of the European Commission [EC] (2003). As a result detailed income data are no longer collected. We therefore restrict the sample to the period 2002 to 2007, for which the survey contains detailed monthly data on spending and income.

The interview unit is the household, defined as people who live in the same main residence and share their expenses. The household is asked to designate the person in the household with the highest income as the *household head*. The household is then asked to provide information for a given month, and consumption, income and all other data relate to this particular month.

The Estonian HBS contains detailed information on twelve different consumption categories, but we follow Pissarides and Weber (1989) and use only the category depicting spending on food, including spending on eating outside the home. They argue that food consumption is the spending category that is most reliably reported by the self-employed.

Households provide detailed listings of different income sources for the month of the interview. The *after-tax* current income of the household is the sum of five different income components, (i.e. income from employment, income from self-employment, property and capital income, transfers, and other income components). The income from self-employment, which we label *business income*, includes income from registered self-employment activities, provision of services, domestic production, intermediation and trade, and other business activities.

The Estonian HBS is unique in that the household is asked to specify not only the *current* monthly after-tax income of the household in the month for which data are collected, but also the typical or *regular* monthly after-tax income. The question about regular income is asked after the detailed data on the current income have been provided. The regular income excludes temporary or extraordinary income components. The properties of the

regular and transitory income components are investigated in detail in Kukk, Kulikov, and Staehr (2015), and their findings suggest that the regular income variable is a useful measure of the permanent income that Pissarides and Weber (1989) suggest as an explanatory variable.

By using permanent rather than current income, it is possible to provide a point estimate of the extent of underreporting instead of a range. If current income is used, it is only possible to provide a range for the extent of underreporting as additional assumptions on the covariation between permanent income and underreporting are needed (Pissarides and Weber 1989). We can use the regular income in the estimations, but as a robustness check we also provide the estimations using the current income measure. Moreover, we include a temporary income measure as a control variable to account for excess sensitivity, which may be misinterpreted as income underreporting. Temporary income is computed as the difference between current income and regular income.

The spending on food is deflated by the HICP consumer price index for food, while the income variables are deflated by the overall HICP index; the resulting real variables are expressed in 2005 prices.

The Estonian HBS makes it possible to identify households prone to underreporting in two different ways. Under the first method a household is seen as prone to underreporting if it reports business income above a given threshold. It is noticeable that households with business income typically also have income from other sources; very few households report having only income from self-employment.

The second method of identification can be derived from a question in the Estonia HBS on the employment status of the household head. If the household head has several income sources, it is instructed to report the employment status for the main income source. The household has the status of self-employed if the household head is "an entrepreneur with employees, including farmer," "a sole proprietor or freelancer," or "employed in a family enterprise, including farmer." It is notable that this means of identification refers only to the employment status of the *household head*; in households with two adults the status of the other adult is not known.

The HBS makes it possible to include a large number of control variables in the estimations. The variables include a dummy indicating the region in which the household resides; the age, gender, nationality and education level of the household head; the number of children under 16; and a dummy variable that takes the value one for households with two income earners. Unfortunately no explicit wealth variables are available in the Estonian

Table 1. Number of Households Using Different Definitions of Self-employment.

		Reported employment status		
		Wage earners	Self-employed	Total
Share of business income in reported income	0%	2,923	162	3,085
	> 0%	2,565	366	2,931
	Total	5,488	528	6,016

Note: The sample includes households with two adults.
Source: Estonian HBS 2002–2007.

HBS; instead we include a dummy taking the value one if the household rents its accommodation and a dummy taking the value one if the household owns real estate in which it does not reside. A list of the variables along with summary statistics is provided in Column (A1.1) in Table A.1 in the Appendix.

All previous studies have restricted the sample to households with two adults. We also focus on households with two adults as this facilitates comparisons with the previous studies and ensures that a large number of observations are available. However, we examine the robustness of the results by repeating the estimations using a sample of households with only one adult.

Table 1 shows how the two identification schemes divide the sample of approximately 6,000 households with two adults into different groups. Less than 10 percent of the households report their status as self-employed. Approximately one third of the households with the status of self-employed do not report any business income, while almost half of the households that report to be wage earners earn some business income. In other words, many households receive both wage income and business income; households with status as self-employed comprise only a small part of the households who report receiving business income.

Summary statistics for the four groups are provided in Columns (A1.2)-(A1.5) in Table A.1 in the Appendix. The number of households with no reported business income is broadly the same as the number of households with some reported business income. The share of business income in reported income is very small in many of the households who report earning some business income.

The large number of households with some business income suggests that many households combine employment and self-employment. According to Eamets (2008), a substantial share of employed people with more

than one job are self-employed in their second job. In order to encourage entrepreneurship, the Estonian authorities made efforts in the beginning of the 2000s to simplify the legislation regulating entrepreneurship, including the registration of a company or a business as a sole proprietor (Leetmaa and Võrk 2007).[3]

Table A.1 in the Appendix shows the mean of regular income and food consumption for different groups of households with two adults. The groups exhibit very different shares of food consumption out of regular income. This observation motivates a detailed investigation of the consequences of using different identification methods.

Empirical Model

This section provides a brief description of the Pissarides and Weber (1989) methodology used in the paper; Kukk and Staehr (2014) discuss the modified version that applies when data on regular income are available. The main assumption is that the self-employed, seen as prone to underreporting, and other households only differ in the extent to which the households report their income. Household i reports regular income $Y_i^{\text{rep\_reg}}$, where regular income is the income when temporary income components are excluded. The log of total permanent income can be expressed as:

$$\log Y_i = \log Y_i^{\text{rep\_reg}} + \log \kappa_i \tag{1}$$

To obtain the total permanent income Y_i, the reported regular income $Y_i^{\text{rep\_reg}}$ has to be multiplied by a factor κ_i; income underreporting is present if $\kappa_i \geq 1$. Self-employed households are generally assumed to report a smaller fraction of their total regular income than households in the comparison group.[4] It is assumed that κ_i is a random variable with a log-normal distribution so $\log \kappa_i$ is a random variable with mean μ_κ and variance $\sigma_v^2 > 0$. The average underreporting factor can be found as:

$$\kappa = \exp(\mu_\kappa + \frac{1}{2}\sigma_v^2) \tag{2}$$

The first step is to estimate a food consumption function or Engel curve. As consumption decisions are determined by the total permanent income but only reported regular income is observed, equation (1) is used to replace the total income in the Engel curve:

$$\log C_i = \alpha + \beta \log Y_i^{\text{rep\_reg}} + \beta \log \kappa_i + X_i'\delta + \varepsilon_i \tag{3}$$

The variable $\log C_i$ is the log real food consumption and the variable $\log Y_i^{\text{rep\_reg}}$ is the log reported regular real income. The control variables in the column vector X_i seek to account for household heterogeneity, which affects food consumption. The error term is labelled ε_i.

When the self-employed households, i.e. households prone to underreporting, and a relevant comparison group are pinned down by the chosen identification method, the following food consumption model can be estimated:

$$\log C_i = \alpha + \beta \log Y_i^{\text{rep\_reg}} + \gamma D_i + X_i'\delta + \varepsilon_i \tag{4}$$

The dummy variable D_i takes the value one for any household identified as self-employed. The coefficient γ equals $\beta\mu_\kappa$. The computation of κ in equation (2) uses the estimated coefficients γ and β and the variance of the underreporting factor, σ_v^2, which is not known.

For the computation of σ_v^2, income regressions are estimated for the two subsamples separately:

$$\log Y_i^{\text{rep\_reg}} = \alpha + Z_i'\varphi + \omega_i \tag{5}$$

where Z_i is a column vector of variables determining the reported regular income and ω_i is unexplained variation. The variance of the error term of the income regression in equation (5) is denoted σ_{SE}^2 for the group of self-employed households and σ_0^2 for the comparison group. It is assumed that the source of any difference between the variances of the error terms in the income regressions of the different subgroups stems from the variance in the log underreporting factor.[5] The mean underreporting factor can then be found as:

$$\kappa = \exp\left(\frac{\gamma}{\beta} + \frac{1}{2}(\sigma_{SE}^2 - \sigma_0^2)\right) \tag{6}$$

The mean underreporting factor κ is the figure by which the regular income reported by the self-employed households must be multiplied in order for these households to attain on average the same propensity for food consumption as households in the comparison group. A mean underreporting factor of $\kappa > 1$ entails that the self-employed households underreport their income to a larger extent than the comparison group do.

Another way to express the extent of underreporting is the underreporting share s, or:

$$s = \frac{(\kappa - 1)}{\kappa}. \tag{7}$$

The underreporting share s is the mean share of unreported income in total (reported and unreported) income relative to the comparison group. If the comparison group does not underreport income, the underreporting share in equation (7) can be seen as the share of total income not reported.

The standard errors of κ and s cannot be computed directly because the standard errors of σ^2_{SE} and σ^2_0 are not known. The standard errors can, however, be obtained by bootstrapping the sequence of regressions used to estimate κ and s. First, the income regressions are estimated separately for the group prone to underreporting and for the comparison group, from which the standard errors of σ^2_{SE} and σ^2_0 are recouped. Second, the consumption model is estimated. Finally, the underreporting factor κ and the share of unreported income s are estimated. The full sequence is replicated 400 times, providing the distribution of the estimated κ and s from which the standard deviation of the parameters is calculated.

Underreporting Results Using Different Identification Schemes

We start the empirical analysis by examining the results when each of the identification methods is applied separately. The first method is the straightforward division of the full sample into households where the household head is reported to be self-employed and households where the household head is reported to be a wage earner. The second method is the division of the sample by the reported business income of the household.

The second method gives rise to a minor complication as the share of business income is a continuous variable that is turned into a binary variable by using a cut-off value. There is no ex ante preferable cut-off value. Pissarides and Weber (1989) use a cut-off value of 25 percent, so that households with business income of over 25 percent of their reported income are considered as self-employed and prone to underreporting, and the rest are considered as wage earners, including those who report business income of less than 25 percent of their total income. Shuetze (2002) uses a cut-off point of 30 percent; a cut-off at zero is not used in any studies. Kukk and Staehr (2014) investigate the income underreporting of households with different shares of business income in Estonia and show that households with even a small share of business income underreport income, though the share is relatively small. The study uses a threshold of 20 percent for households with business income and excludes from the comparison group households with a smaller share of business income. As we are comparing the results of two identification methods, we prefer not to exclude any households from

the sample when using business income as an indicator and we therefore use a cut-off point of zero when identifying the households prone to underreporting based on their reported business income.

We estimate the consumption model in equation (4) on monthly data from January 2002 to December 2007. The coefficients α, β, and γ are to be estimated along with those in the vector δ. The control variables in the column vector X_i include the log temporary real income, the demeaned age of the household head and its squared value, the number of children aged below 16, a dummy for two income earners in the household, a dummy for renting of the main residence, and a dummy for ownership of real estate beyond the main residence. Finally, 71 monthly dummies are included as control variables to account for time-dependent aggregate shocks.

We use cross-sectional data and can therefore not control for unobserved heterogeneity so the income variable $\log Y_i^{\text{rep\_reg}}$ is likely to be correlated with the error term, which calls for the variable to be instrumented. We estimate several different model specifications and for comparability the same set of instruments have been used in all specifications, i.e. the education level, gender and nationality of the household head and regional dummies for the log regular income $\log Y_i^{\text{rep\_reg}}$. According to the standard Mincer model, education level is an important determinant of employment income (Mincer 1974). Empirical studies for Estonia have shown significant income gaps between men and women and between people of different ethnicities (Leping and Toomet 2008). Regional NUTS-3 dummies are used to capture the regional labour market situation (Blanchflower and Oswald 1995; Verkulevičiūtė-Kriukienė 2015).[6]

The results of the estimation of the consumption model in equation (4) are shown in Table A.2 in the Appendix for the two identification methods considered. The estimated coefficients are very similar across the two cases; only the coefficient of the dummy indicating underreporting households varies substantially, suggesting different underreporting factors.

Table 2 shows the average underreporting factor κ defined in equation (6) and the underreporting share s defined in equation (7), both relative to the comparison group. The results are different when different methods are used to identify the households prone to underreporting. Note that the way in which the group of households prone to underreporting is identified also determines the group taken to be the comparison group. Column (2.1) shows the results when the identification is based on reported employment status. The reported income of households reporting their status as self-employed must be multiplied by the underreporting factor $\kappa = 1.386$ for

Table 2. Estimates of Income Underreporting Relative to the Comparison Group Using Different Identification Methods.

	(2.1) Reported employment status[a]	(2.2) Reported share of business income[b]
Underreporting factor κ	1.386 (0.068)	1.747 (0.057)
Underreporting share s	0.279 (0.036)	0.428 (0.018)
Observations	6,016	6,016

Note: The sample includes households with two adults. Calculations using the consumption estimations in Columns (A2.1) and (A2.2) in Table A.2 are in the Appendix. Bootstrapped standard errors are reported in parentheses.

[a]The underreporting group consists of household in which the head is reported to be self-employed, while the comparison group consists of households where the head is reported to be a wage earner.

[b]The underreporting group consists of households that report some business income, while the comparison group consists of households that do not report any business income.

these households to have the same propensity for food consumption as households with the status of wage earners. It follows that the households with self-employment status leave approximately 28 percent of their total income unreported. These measures of underreporting are, as whenever the Pissarides and Weber (1989) approach is used, relative to the underreporting in the reference group, in this case the households with the status of wage earners.

Column (2.2) shows the underreporting results when the households prone to underreporting are taken to be those that report business income above 0. The underreporting factor $\kappa = 1.747$ is substantially larger than before; the reported income of households with any business income must be multiplied by 1.7 to attain the same consumption propensity as for households with no business income. This implies that the households with business income above 0 leave around 43 percent of their total income unreported compared to households with no business income.[7]

A higher underreporting factor for households with a higher share of business income is an expected outcome if the underreporting pertains mainly to business income. The estimated underreporting factor shows how much income is unreported from total (reported and unreported) income, and it is reasonable to assume that the underreporting factor should be higher for households with a larger share of business income. Table S.1 in the online supplementary materials provides the estimates of the underreporting factor for households with different shares of business income,

confirming that a higher share of business income concurs with a higher share of unreported income.[8] When households with some business income are excluded from the comparison group of wage earners, the underreporting factor is 2.6 and the underreporting share is 62 percent.

This paper focuses on the consequences of the use of different identification methods. Kukk and Staehr (2014) discuss possible reasons for such extensive underreporting in Estonia and point in particular to the possibility of private spending being reported as deductible business expenses. Paulus (2015b) similarly finds very substantial underreporting using Estonian data and provides some explanation for the results.

The results in Table 2 taken in combination paint an interesting picture. The two identification methods lead to markedly different results. The share of unreported income is estimated to be substantially larger when the households prone to underreporting are identified using the share of business income than when they are identified using their employment status. The differences are even more striking when the standard approach of the other studies is applied and a cut-off point at some positive value is used for business income.

As a robustness test, we have run the estimations using current income rather than regular income; in this case we can only provide the range of the estimated underreporting factor. The estimated coefficients of the consumption model and the estimated underreporting factors are given in Table S.2 in the online supplementary materials. The coefficients of the consumption model are very similar to those of the baseline model, which are given in Table A.2. The range of the estimated underreporting factor is quite large, as in other studies, but the differences between the results using the different identification methods are evident.

These findings show the importance of the choice of identification method. The results of studies using the Pissarides and Weber (1989) methodology must be interpreted in light of their chosen identification method. Using only the reported employment status we attain the results shown in Column (2.1), where underreporting is estimated at around 28 percent of total income relative to the comparison group. This result is in the same range as the results for the United States, Finland, and Sweden, which identify the self-employed by their reported employment status, cf. Engström and Holmlund (2009), Hurst, Li, and Pugsley (2014), and Johansson (2005). Our results based on Estonian data suggest, however, that the share of underreporting is substantially higher when the households prone to underreporting are identified from their reported business income.

Underreporting Results Combining Different Identification Schemes

To obtain a better understanding of the factors behind the results of the different identification methods in Section 5, we distinguish between the different groups or sub-samples shown in Table 1. This entails combining information on the employment status and the share of business income of the households. The objective is to examine whether combining the two different criteria leads to a better identification of the households that are prone to underreporting.

The Baseline Estimations

We take as the comparison group the households of wage earners in which the household head is reported to be a wage earner *and* which do not report any business income. We consider three groups of households prone to underreporting, indexed by g and differentiated by the way the underreporting households are identified. The first group ($g = 1$) contains households in which the household head is a wage earner *and* the household reports some business income. The second group ($g = 2$) contains households in which the head is reported to be self-employed *and* the household reports some business income. The third group ($g = 3$) consist of households where the head is self-employed *and* no business income is reported. The number of households is 2,565 in the first group, 366 in the second group and 162 in the third group, as also reported in Table 1.

The more homogeneous groups defined above can easily be mapped onto the groups used in the estimations presented in Table 2. In Column (2.1) the second and third groups are in the group prone to underreporting, whereas the first group is included in the comparison group together with the group of wage earners without business income. In Column (2.2) however, the first and the second groups comprise the underreporting group while the third group and the wage earners are in the comparison group.

Using the more fine-grained division of the sample leads to the following consumption model:

$$\log C_i = \alpha + \beta \log Y_i^{\text{rep\_reg}} + \gamma_1 D_{1i} + \gamma_2 D_{2i} + \gamma_3 D_{3i} + X_i'\delta + \varepsilon_i \qquad (8)$$

where the self-employment dummies D_1, D_2, and D_3 are used for the three groups of underreporting households identified using different identification methods. The estimated coefficients γ_1, γ_2, and γ_3 of the dummies are

used to compute the underreporting factors $\kappa_g = \kappa_1$, κ_2 and κ_3 and the underreporting shares $s_g = s_1$, s_2 and s_3 for the three groups of households.

The estimation results for the consumption model are reported in Column (A2.3) in Table A.2 in the Appendix. Column (3.1) in Table 3 shows the computed underreporting factor κ_g for each group and Column (3.2) shows the corresponding underreporting shares s_g, both being relative to the comparison group of wage earners without any business income.

The estimations where the underreporting households are identified through their business income provide higher underreporting results when the household head has the status of self-employed than when the head is a wage earner. The income of wage earners with business income should be multiplied by 1.7 to attain a total income comparable to that of the comparison group of wage earners who do not report any business income. The income of households in which the household head has the status of self-employed and which have some business income should be multiplied by 2.2. In other words, the underreporting share for households stating themselves to be wage earners is 41 percent and the share for households stated as self-employed is 55 percent, in both cases relative to the comparison group of wage earners without any business income.

Estimations with Matched Samples

Further analysis suggests that the difference between wage earners and the self-employed with business income largely stems from the different distributions of the share of business income in reported income among wage earners and the self-employed. The left panel in Figure 1 shows the distribution of the share of business income for the two groups and it is clear that there are a substantial number of wage earners with a relatively low share of business income. To take account of the different shares of business income across the two groups, a one-to-one matching was carried out using the share of business income. For each household in the group of the self-employed, a household with a similar share of business income was chosen from the group of wage earners.[9] As a result a total of 366 households (rather than 2,565 households) were included in the group of wage earners with business income; the same number of households is in the group of self-employed with business income. The right panel in Figure 1 confirms that the distribution of the share of business income after the matching is comparable in the two samples.

The underreporting results using the matched sample are given in Columns (3.3) and (3.4) in Table 3. The underreporting factor is 2.3 for

Table 3. Estimates of Income Underreporting Combining Different Identification Schemes. Households with Two Adults.

	(3.1)	(3.2)	(3.3)	(3.4)
	Full sample		Matched sample	
	Underreporting factor κ_g	Underreporting share s_g	Underreporting factor κ_g	Underreporting share s_g
Wage earners, share of business income > 0	1.705 (0.056)	0.413 (0.019)	2.295 (0.150)	0.564 (0.030)
Self-employed, share of business income > 0	2.214 (0.137)	0.548 (0.029)	2.148 (0.124)	0.534 (0.028)
Self-employed, share of business income = 0	1.054 (0.071)	0.051 (0.072)	1.026 (0.065)	0.025 (0.070)
Wald-test (H_0: $\kappa_1 = \kappa_2$)	19.19 [0.000]		0.82 [0.364]	
Wald-test (H_0: $\kappa_3 = 1$)	0.58 [0.448]		0.16 [0.693]	
No. of obs.	6,016		3,817	

Note: Calculations use the consumption estimations in Column (A2.4) in Table A.2. The comparison group of the estimations contains households who report that the household head is a wage earner and the household does not earn any business income. In the matched sample the group of wage earners with business income above 0 is matched to the group of the self-employed with business income above 0 based on the share of business income. Bootstrapped standard errors are reported in parentheses. A Wald-test provides the χ^2 statistic for the null hypothesis while the *p*-value is given in square brackets.

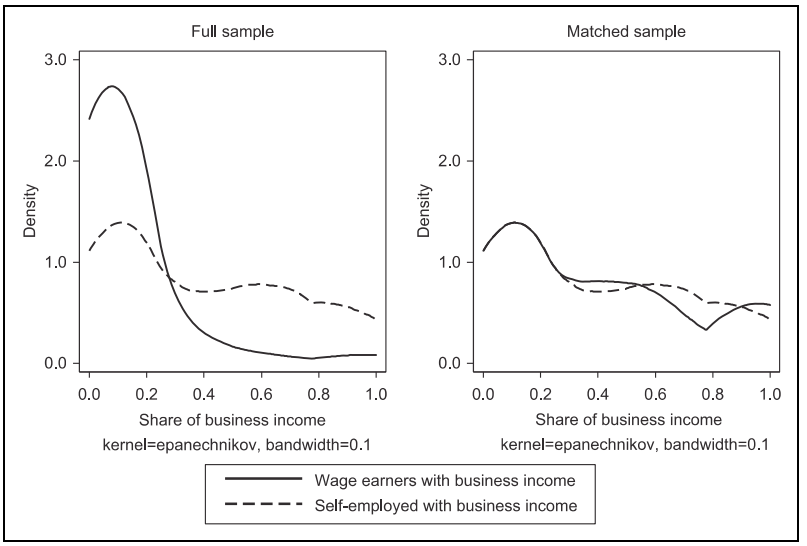


Figure 1. Kernel density estimates.
Note: The full sample includes all wage earners with business income. In the matched sample households of wage earners with business income are matched to households of the self-employed using the share of business income.

households where the head is a wage earner and this is not statistically different from the estimated underreporting factor for households where the head is self-employed. In other words, the difference in the results between wage earners and the self-employed is marginal when the share of business income is taken into account. It can be seen from Columns (A1.3) and (A1.4) in Table A.1 in the Appendix that these households exhibit quite comparable socio-demographic variables, suggesting that the two groups are relatively similar. As seen in Columns (A1.2) and (A1.3), the profile of wage earners without business income is somewhat different from the profile of wage earners with business income, implying that the two groups should not be pooled.

The sample has hitherto comprised only households with two adults who are active in the labour market. Such a sample composition is chosen in all studies using the Pissarides and Weber (1989) method, as the homogeneous sample makes it more likely that any differences in food consumption are driven by underreporting. Our results utilizing the reported employment status may, however, be questioned as we observe the employment status only

for the household head and no information on the employment status of the second adult is available. To ensure that this particularity does not drive the results, we repeat the estimations in equation (8) for households with only one adult so that the employment status is known for all the adults in the household. The results for households with one adult are provided in Table S.3 in the online supplementary materials, and they show that the results are comparable to those for households with two adults.

Discussion of the Results

It follows from Column (2.1) that if the identification is based entirely on reported employment status, the estimated share of underreporting relative to the comparison group is around 28 percent of total income. However, we obtain quite different results when we split those with self-employment status into two groups. Column (3.2) shows that the share of underreporting by households whose status is self-employed varies markedly depending on the reported business income; the underreporting is estimated to be around 55 percent if business income is reported, but essentially zero if no business income is reported.

The results in Table 3 explain those in Table 2. The estimated underreporting share of 28 percent when there is no conditioning on reported business income, shown in Column (2.1) in Table 2, derives from a mixture of different effects. First, it follows from Table 1 that 47 percent of the comparison group of wage earners in Column (2.1) in Table 2 earn some business income, while Table 3 shows that wage earners with business income underreport 41 percent of their total income. If it is assumed that wage earners without business income report their income truthfully, the average underreporting in the comparison group would be $0.47 \cdot 0.41 \approx 19$ percent. Second, as the group of self-employed in Column (2.1) in Table 2 includes the self-employed without business income, the estimations essentially give the weighted average of the estimated underreporting share for the two sub-samples, i.e. the self-employed with and without business income, which is $0.69 \cdot 0.55 + 0.31 \cdot 0.05 \approx 40$ percent. Column (2.1) shows the *additional* underreporting of the self-employed relative to the underreporting of wage earners, or $(0.4 - 0.19)/(1 - 0.19) \approx 25$ percent, which is very close to the estimated underreporting share of 28 percent in Column (2.1) in Table 2.[10]

When the wage earners without business income are used as a comparison group and other groups are taken to be prone to underreporting, the estimations exhibit the underreporting of different groups relative to the

same comparison group. The estimations in Table 3 reveal that the estimated underreporting shares s_1 and s_2 are similar while s_3 is very small. In other words, the reported employment status of the household does not matter if the share of reported business income is taken into account. Conditional on the reported share of business income, households with the status of self-employed and households of wage earners appear to underreport income to the same extent.

The results in Table 3 are robust and also emerge when we use OLS estimations instead of IV estimations, or when non-durable consumption is used instead of food consumption; see Tables S.4 and S.5 in the online supplementary materials. Finally, additional robustness tests have been carried out using smaller sets of instruments (not reported), but the estimation results are very similar, indicating that the results are not sensitive to the choice of instruments.

The underreporting factor for households with the status of self-employed but no reported business income is not statistically different from 1; there is no sign of income underreporting for this group relative to the comparison group of wage earners without business income. It follows from Table 1 that approximately one third of the households who have the status of self-employed do not report any business income. A possible reason for the absence of business income in households that report their status as self-employed could be that they are fully concealing the self-employment income, but this hypothesis is ruled out by the finding of no underreporting. It is therefore likely that these households indeed have no income because of temporary inactivity or adverse conditions affecting their business.

Final Comments

Limited third-party reporting implies that the self-employed have ample opportunities to underreport their income. The methodology of Pissarides and Weber (1989) is an ingenious means of assessing the extent of possible income underreporting by this group. The studies applying the P&W methodology use two different ways to identify the households prone to underreporting, i.e. the reported employment status and the share of reported business income in reported income. This paper is the first to examine in detail the effect of the choice of method for identifying underreporting households when the Pissarides and Weber (1989) methodology is used. The estimations are implemented using survey data, and the results may be seen as an indication of possible underreporting to the tax authorities (Hurst, Li, and Pugsley 2014; Pissarides and Weber 1989).

The Estonian Household Budget Survey contains detailed information on both consumption and income measures along with the reported employment status, and this makes it possible to use either of the two main identification schemes. The share of unreported income is estimated to be around 28 percent of total (reported and unreported) income if the employment status is used for identifying households prone to underreporting, while it is 43 percent if households prone to underreporting are identified using reported business income. The unconditional estimates are based on sample splits along only one dimension as is customarily done in the literature using the Pissarides and Weber (1989) method.

Additional insights emerge when the sample is split along both dimensions, i.e. business income and reported employment status. When the distributions of the business income shares for wage earners and self-employed are matched, the extent of underreporting turns out to be very similar for the reported self-employed with business income and for reported wage earners with business income. These results are robust to different household compositions, including households with one adult. The upshot is that the identification method is of material importance, as reporting of business income appears to be a good indicator of potential underreporting; a substantial share of households prone to income underreporting are overlooked if only households that report themselves to be self-employed are considered prone to underreporting.

Table A1. Descriptive Sample Statistics.

	(A1.1)	(A1.2)	(A1.3)	(A1.4)	(A1.5)
	Full sample	Reported wage earners, business income = 0	Reported wage earners, business income > 0	Reported self-employed, business income > 0	Reported self-employed, business income = 0
Sample size	6,016	2,923	2,565	366	162
Log (Food consumption)	7.72	7.55	7.88	7.98	7.68
Log (Current income)	9.07	9.06	9.08	9.01	9.20
Log (Regular income)	8.96	8.98	8.92	8.93	9.20
Share of business income	0.08	0.00	0.13	0.40	0.00
Age	43.47	42.11	44.77	46.21	41.12
Female	0.38	0.40	0.40	0.20	0.14
Non-Estonian	0.25	0.38	0.13	0.09	0.30
Education level 1	0.05	0.05	0.05	0.06	0.01
Education level 2	0.52	0.51	0.53	0.46	0.46
Education level 3	0.43	0.44	0.42	0.48	0.53
Number of children	0.74	0.66	0.80	0.87	0.92
Region North	0.29	0.41	0.16	0.15	0.45
Region Northeast	0.11	0.16	0.06	0.02	0.04
Region Central	0.13	0.10	0.17	0.17	0.07
Region West	0.17	0.12	0.20	0.27	0.19
Region South	0.31	0.21	0.41	0.39	0.25
Two income earners	0.55	0.56	0.53	0.57	0.54
Owning second property	0.09	0.09	0.09	0.07	0.10
Renting main residence	0.07	0.09	0.06	0.03	0.10

Notes: The table gives the mean values of the variables; socio-demographic variables are for the household head while the economic variables are for the household.

Table A2. Full Estimations of IV Regression Model. Dependent Variable: Food Consumption.

	(A2.1)	(A2.2)	(A2.3)	(A2.4)
	Reported employment status[a]	Reported share of business income[b]	Two adults Full sample	Two adults Matched sample
Log regular income (β)	0.567*** (0.032)	0.596*** (0.031)	0.588*** (0.031)	0.613*** (0.039)
Self-employment dummy (γ)	0.152 (0.024)			
Business income dummy (γ)		0.320*** (0.012)		
Dummy for wage earners, business income > 0			0.303*** (0.012)	0.456*** (0.032)
Dummy for the self-employed, business income > 0			0.429*** (0.028)	0.429*** (0.028)
Dummy for the self-employed, business income = 0			−0.005 (0.041)	−0.022* (0.041)
Log temporary income	0.213*** (0.018)	0.183*** (0.016)	0.184*** (0.016)	0.157*** (0.017)
Age – mean age	0.004*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)
(Age – mean age)^2/10	−0.002*** (0.000)	−0.002*** (0.000)	−0.002*** (0.000)	−0.002*** (0.000)
Number of children	0.110*** (0.008)	0.087*** (0.008)	0.086*** (0.008)	0.092*** (0.010)
Two income earners	−0.053*** (0.018)	−0.059*** (0.017)	−0.058*** (0.017)	−0.057*** (0.022)
Owning second real estate	−0.074*** (0.020)	−0.062*** (0.019)	−0.059*** (0.019)	−0.069*** (0.025)
Renting main residence	−0.045* (0.025)	−0.030 (0.024)	−0.028 (0.024)	0.045 (0.028)
Constant	2.729*** (0.273)	2.322*** (0.263)	2.396*** (0.262)	2.230*** (0.333)

(continued)

Table A2. (continued)

	(A2.1)	(A2.2)	(A2.3)	(A2.4)
	Reported employment status[a]	Reported share of business income[b]	Two adults Full sample	Two adults Matched sample
R^2	0.228	0.296	0.302	0.287
R^2 of the first regression	0.424	0.425	0.428	0.435
F-stat of the first regression [*p*-value]	54.53 [0.000]	54.53 [0.000]	54.84 [0.000]	26.81 [0.000]
No. of obs.	6,016	6,016	6,016	3,817

Notes: IV estimations are reported, where 71 monthly dummies are included in the estimations but are not shown in the table. In the matched sample the group of wage earners with business income above 0 is matched to the group of self-employed with business income above 0 using the share of business income. Robust standard errors are reported in round parentheses. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% levels respectively. *F*-stat is the test statistic for the Wald test of overall goodness of the fit of the first regression.

[a]Self-employed vs. wage earners.

[b]Business income > 0 vs. business income = 0.

Acknowledgments

We would like to thank three anonymous referees and the editor of the *Public Finance Review* as well as Alari Paulus, Tairi Rõõm, and participants in the Shadow2013 conference and in a seminar at Tallinn University of Technology for valuable comments. We are grateful to Arvo Valtin from Statistics Estonia and Dmitry Kulikov from Eesti Pank for making the database available. The views expressed are those of the authors and not necessarily those of Eesti Pank or other parts of the Eurosystem.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Support from Estonian Target Financing grant no. SF0140059s12 and Base Financing grant no. B45/2015 is acknowledged.

Notes

1. The Living Cost and Food Survey in the UK collects data on weekly spending with diaries while requiring the households to report only their usual weekly income without detailing the source. The Panel Survey of Income Dynamics (PSID) contains detailed income data but the survey collects food consumption data at a rather general level as no diary is used.
2. An example of this is the US Consumer Expenditure Survey (CEX), in which households report annual income for the previous year, while consumption data are collected for the current quarter.
3. The application can be submitted to the registrar electronically without notary involvement and at low cost, and the registration is typically granted only a few hours after it is submitted. In addition, it is relatively simple to file tax returns through the website of the tax authorities.
4. The assumption of no income underreporting has customarily been used for the comparison group. However, the current specification makes it possible to interpret the results in terms of income underreporting in addition to the underreporting by the comparison group. This interpretation is useful if wage earners exhibit substantial income underreporting. Studies suggest, however, that underreporting of wage income may be relatively modest in Estonia (Kriz et al. 2008; Williams 2009).

5. After controlling for household specific economic and socio-demographic characteristics, the variance of the residual in the regression of regular income in equation (5) contains unexplained variation in regular income and deviations of total from reported regular income. When the unexplained variation in regular income has the same variance for the underreporting group and the comparison group, mainly stemming from omitted variables, then the difference between the residual income variance of the two groups derives from the variance σ_v^2 . This assumption is used by Pissarides and Weber (1989) as well as by other authors applying the method.
6. In the dataset the education level, gender, and nationality are valid instruments in all specifications while there is some doubt about the regional dummies in some specifications. There is no apparent reason for regional dummies to affect food consumption in Estonia as prices are very similar across the country due to the small size of the market and relatively short distances between the regions. There are, however, substantial regional disparities in income and employment opportunities in Estonia which cannot be accounted for by variables for education, gender and ethnicity (Verkulevičiūtė-Kriukienė 2015). We obtain qualitatively similar results when using a smaller number of instruments (not shown; the estimations are available upon request).
7. A cut-off point at some positive value for business income is typically used in studies exploiting the P&W methodology. For comparison we carried out estimations with a cut-off point of 20 percent where the comparison group consists of households with no business income or with business income of up to 20 percent of their reported income. The underreporting factor is 2.3 and the underreporting share is around 56 percent for the households with business income above 20 percent (not shown).
8. Kukk and Staehr (2014) discuss in detail the importance of the threshold value of the share of business income in estimating underreporting of income. The present paper does not consider different thresholds for the share of business income as the main focus is on the comparison of different identification schemes. As the employment status can take only binary values, the business income measure is similarly kept as a binary variable.
9. For selection of the matching pairs we use single nearest neighbour matching with a common support and without replacement.
10. This can also be expressed in algebra. If in the comparison group a fraction λ_0 are pure wage earners with 0 underreporting and a fraction $\lambda_1 = 1 - \lambda_0$ underreport income by the share of s_1, the average share of underreported income in the comparison group is $\lambda_1 s_1$. The underreporting group consists of a fraction λ_2 that underreports s_2 and a fraction $\lambda_3 = 1 - \lambda_2$ whose underreporting share is s_3, implying that on average $\lambda_2 s_2 + (1 - \lambda_2)s_3$ of the income is not reported in this

group. The estimated underreported share in Column (2.1) can be calculated as $(\lambda_2 s_2 + (1 - \lambda_2)s_3 - \lambda_1 s_1)/(1 - \lambda_1 s_1)$.

Supplemental Material

The online [appendices/data supplements/etc] are available at http://pfr.sagepub.com/supplemental.

References

Ali, Mukhtar M., H. Wayne Cecil, and James A. Knoblett. 2001. "The Effects of Tax Rates and Enforcement Policies on Taxpayer Compliance: A Study of Self-employed Taxpayers." *Atlantic Economic Journal* 29:186–202.

Alm, James, Roy Bahl, and Matthew N. Murray. 1991. "Tax Base Erosion in Developing Countries." *Economic Development and Cultural Change* 39:849–72.

Alm, James, John Deskins, and Michael McKee. 2009. "Do Individuals Comply on Income Not Reported by Their Employer?" *Public Finance Review* 37:120–41.

Blanchflower, David G., and Andrew J. Oswald. 1995. "An Introduction to the Wage Curve." *Journal of Economic Perspectives* 9:153–67.

Bruce, Donald. 2000. "Effects of the United States Tax System on Transitions into Self-employment." *Labour Economics* 7:545–74.

Eamets, Raul. 2008. *Decent Work Country Report—Estonia*. Geneva, Switzerland: International Labour Organisation. Accessed November 3, 2015. http://www.ilo.org/.

EC (European Commission). 2003. *Household Budget Surveys in the EU. Methodology and recommendation for harmonisation–2003*. Luxemburg: European Communities.

Engström, Per, and Bertil Holmlund. 2009. "Tax Evasion and Self-employment in a High-tax Country: Evidence from Sweden." *Applied Economics* 41:2419–30.

Fabrizio, Stefania, and Ashoka Mody. 2010. "Breaking the Impediments to Budgetary Reforms: Evidence from Europe." *Economics and Politics* 22:362–91.

Gibson, John, Bonggeun Kim, and Chul Chung. 2009. "Using Panel Data to Exactly Estimate Income Under-reporting of the Self Employed." Labor Economics Working Papers, no. 22996, East Asian Bureau of Economic Research, Canberra, Australia.

Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110:458–74.

Hurst, Erik, Geng Li, and Benjamin Pugsley. 2014. "Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self Employed." *Review of Economics and Statistics* 96:19–33.

Johansson, Edvard. 2005. "An Estimate of Self Employment Income Underreporting in Finland." *Nordic Journal of Political Economy* 31:99–109.

Kriz, Kenneth, A., Jaanika Meriküll, Alari Paulus, and Karsten Staehr. 2008. "Why Do Individuals Evade Payroll and Income Taxation in Estonia?" In *Shadow Economy, Corruption and Governance*, edited by Michael Pickhardt and Edward Shinnick, 240–64. Cheltenham, UK: Edward Elgar Publishing.

Kukk, Merike, Dmitry Kulikov, and Karsten Staehr. 2015. "Estimating Consumption Responses to Income Shocks of Different Persistence Using Self-reported Income Measures." *Review of Income and Wealth*, doi: 10.1111/roiw.12163.

Kukk, Merike, and Karsten Staehr. 2014. "Income Underreporting by Households with Business Income." *Evidence from Estonia, Post-communist Economies* 26:257–76.

Leetmaa, Reelika, and Andres Võrk. 2007. *Article on Undeclared Work from SYSDEM Correspondent: Estonia*. Brussels, Belgium: European Employment Policy Observatory. Accessed November 3, 2015. http://www.eu-employment-observatory.net/.

Leping, Kristian-Olari, and Ott Toomet. 2008. "Emerging Ethnic Wage Gap: Estonia during Political and Economic Transition." *Journal of Comparative Economics* 36:599–619.

Meriküll, Jaanika, and Karsten Staehr. 2010. "Unreported Employment and Envelope Wages in Mid-transition: Comparing Developments and Causes in the Baltic Countries." *Comparative Economic Studies* 52:637–70.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings*. New York: NBER Press.

Paulus, Alari. 2015a. "Tax Evasion and Measurement Error: An Econometric Analysis of Income Survey Data Linked with Tax Records." ISER Working Paper Series, no. 2015-10, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Paulus, Alari. 2015b. "Income Underreporting Based on Income-expenditure Gaps: Survey vs Tax Records." ISER Working Paper Series, no. 2015-15, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Pissarides, Cristopher, and Guglielmo Weber. 1989. "An Expenditure Based Estimate of Britain's Black Economy." *Journal of Public Economics* 39:17–32.

Schneider, Friedrich, and Colin C. Williams. 2013. *The Shadow Economy*. London, UK: Institute of Economic Affairs.

Schuetze, Herb. 2002. "Profiles of Tax Noncompliance among the Self Employed in Canada: 1969–1992." *Canadian Public Policy* 28:219–37.

Soos, Piroska. 1990. "Self-employed Evasion and Withholding: A Comparative Study and Analysis of the Issues." *U.C. Davis Law Review* 24:107–94.

Staehr, Karsten. 2004. "The Economic Transition in Estonia. Background, Reforms and Results." In *Contemporary Change in Estonia*, edited by Rindzeviciute Egle, 4:37–67. Huddinge, Sweden: Baltic and East European Studies, Sodertorns hogskola.

Staehr, Karsten. 2009. "Estimates of Employment and Welfare Effects of Personal Labour Income Taxation in a Flat-tax Country: The Case of Estonia." In *Microfoundations of Economic Success: Lessons from Estonia*, edited by Mayes, D., 242–90. Cheltenham, UK: Edward Elgar Publishing.

Tafenau, Egle, Helmut Herwartz, and Friedrich Schneider. 2010. "Regional Estimates of the Shadow Economy in Europe." *International Economic Journal* 24:629–36.

Verkulevičiūtė-Kriukienė, Daiva. 2015. "Labour Income and the Unemployment Rate in the Regions of Lithuania, Latvia and Estonia: Differences and Change in the Period 2004-2013." *Regional Formation and Development Studies* 15: 104–15.

Williams, Colin C. 2008. "Envelope Wages in Central and Eastern Europe and the EU." *Post-communist Economies* 20:363–76.

Williams, Colin C. 2009. "The Commonality of Envelope Wages in Eastern European Economies." *Eastern European Economics* 47:37–52.

Author Biographies

Merike Kukk is research scientist at the Department of Finance and Economics at Tallinn University of Technology, Estonia. Her research topics cover household finance and consumption, tax evasion, and fiscal policy. She has consulted commercial banks and has written analyses about household behaviour in the Baltic States. She has also worked for Eesti Pank, the central bank of Estonia.

Karsten Staehr is professor of international and public finance at the Department of Finance and Economics at Tallinn University of Technology, Estonia, and a part-time research supervisor at Eesti Pank, the central bank of Estonia. He holds a master's degree from the Massachusetts Institute of Technology and master's and Ph.D. degrees from the University of Copenhagen. He undertakes research and policy analysis within public economics, monetary economics and transition studies.

Supplementary materials

Identification of Households Prone to Income Underreporting: Employment Status or Reported Business Income?

Public Finance Review

Merike Kukk and Karsten Staehr

Table S.1. Estimates of income underreporting for households with different shares of business income. Dependent variable: food consumption

	(A3.1)	**(A3.2)**	**(A3.3)**
	Estimated coefficients	**Underreporting factor κ_g**	**Underreporting share s_g**
Log regular income	0.625*** (0.031)	..	..
Dummy for business income < 5%	0.131*** (0.016)	1.233 (0.036)	0.189 (0.024)
Dummy for business income 5–10%	0.290*** (0.017)	1.561 (0.055)	0.359 (0.026)
Dummy for business income 10–20%	0.444*** (0.019)	2.005 (0.081)	0.501 (0.020)
Dummy for business income 20–50%	0.581*** (0.039)	2.380 (0.131)	0.580 (0.024)
Dummy for business income ≥ 50%	0.546*** (0.028)	2.988 (0.230)	0.665 (0.030)
R^2	0.331		
No. of obs.	6016		

Notes: Column (A3.1) shows IV estimations with GMM estimator. Education level, nationality, gender and five regional dummies are used as instruments. Log temporary income, age and age squared of the household head, the number of children, two income earners, ownership of a second real estate property, renting the main residence, and 71 monthly time fixed effects are included in the estimations, but the results are not shown in the table. Robust standard errors are reported in round parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% levels respectively. Columns (A3.2) and (A3.3) show underreporting results using the consumption estimations in Column (A3.1). Bootstrapped standard errors are reported in parentheses below the parameter estimates.

Table S.2. Estimates of income underreporting from different identification methods using current income variable. Dependent variable: food consumption

	(A4.1)	**(A4.2)**
	Reported employment status[a]	**Reported share of business income**[b]
A. Consumption model		
Log current income (β)	0.545*** (0.032)	0.553*** (0.031)
Self-employment dummy (γ)	0.165*** (0.027)	…
Business income dummy (γ)	…	0.287*** (0.013)
Age – mean age	0.004*** (0.001)	0.002*** (0.001)
(Age – mean age)^2/10	-0.002*** (0.000)	-0.002*** (0.000)
Number of children	0.103*** (0.009)	0.083*** (0.008)
Two income earners	-0.043** (0.018)	-0.039** (0.018)
Owning second real estate	-0.066*** (0.022)	-0.052** (0.021)
Renting main residence	-0.053** (0.027)	-0.040 (0.026)
Constant	2.882*** (0.273)	2.659*** (0.266)
R^2	0.124	0.178
R^2 of the first regression	0.252	0.251
F-stat of the first regression [p-value]	31.84 [0.000]	31.70 [0.000]

		(A4.1)	(A4.2)
		Reported employment status[a]	**Reported share of business income**[b]
B. Underreporting			
Underreporting factor κ	Lower bound	1.123 (0.071)	1.488 (0.062)
	Upper bound	1.631 (0.122)	1.901 (0.091)
Underreporting share *s*	Lower bound	0.109 (0.050)	0.328 (0.027)
	Upper bound	0.387 (0.050)	0.474 (0.025)
No. of obs.		6016	6016

[a] Self-employed vs. wage earners.
[b] Business income > 0 vs. business income = 0.

Note: Sample of households with two adults. Panel A: IV estimations with GMM estimator. 71 monthly dummies are included in the estimations but are not shown in the table. In the matched sample the group of wage earners with business income above 0 is matched to the group of self-employed with business income above 0 based on the share of business income. Robust standard errors are reported in round parentheses below the coefficient estimates. Superscripts ***, ** and * indicate that the coefficient is statistically different from 0 at the 1%, 5% and 10% levels respectively. Panel B. Calculations using the consumption estimations in Part A. Bootstrapped standard errors are reported in parentheses below the parameter estimates.

Table S.3. Estimates of income underreporting relative to wage earners without any business income. Households with one adult

	(S1.1)	(S1.2)	(S1.3)	(S1.4)
	Full sample		**Matched sample**	
	Underreporting factor κ_g	**Underreporting share s_g**	**Underreporting factor κ_g**	**Underreporting share s_g**
Wage earners, share of business income > 0	1.939 (0.168)	0.484 (0.042)	2.686 (0.425)	0.628 (0.061)
Self-employed, share of business income > 0	1.983 (0.289)	0.496 (0.103)	1.974 (0.298)	0.493 (0.106)
Self-employed, share of business income = 0	1.271 (0.258)	0.213 (0.149)	1.253 (0.251)	0.202 (0.151)
Wald-test (H_0: $\kappa_1 = \kappa_2$)	0.03 [0.871]	..	2.82 [0.093]	..
Wald-test (H_0: $\kappa_3 = 1$)	1.10 [0.294]	..	1.01 [0.314]	..
No. of obs.	2173		1626	

Note: The comparison group of the estimations contains households who report that the household head is a wage earner and the household does not earn any business income. In the matched sample the group of wage earners with business income above 0 is matched to the group of self-employed with business income above 0 based on the share of business income. Bootstrapped standard errors are reported in parentheses below the parameter estimates. A Wald-test provides the χ^2 statistic for the null hypothesis while the *p*-value is given in square brackets below.

Table S.4. OLS estimates of income underreporting relative to the comparison group. Households with two adults. Dependent variable: food consumption

	(S2.1)	**(S2.2)**	**(S2.3)**	**(S2.4)**
	Full sample		**Matched sample**	
	Underreporting factor κ_g	**Underreporting factor s_g**	**Underreporting factor κ_g**	**Underreporting factor s_g**
Wage earners, share of business income > 0	2.258 (0.116)	0.557 (0.022)	3.286 (0.371)	0.695 (0.035)
Self-employed, share of business income > 0	3.182 (0.281)	0.686 (0.029)	3.090 (0.268)	0.676 (0.029)
Self-employed, share of business income = 0	1.183 (0.119)	0.155 (0.095)	1.149 (0.113)	0.130 (0.097)
Wald-test (H_0: $\kappa_1 = \kappa_2$)	14.76 [0.000]	..	0.33 [0.563]	..
Wald-test (H_0: $\kappa_3 = 1$)	2.38 [0.123]	..	1.76 [0.185]	..
No. of obs.	6016		3817	

Note: The comparison group of the estimations contains households who report that the household head is a wage earner and the household does not earn any business income. In the matched sample the group of wage earners with business income above 0 is matched to the group of self-employed with business income above 0 based on the share of business income. Bootstrapped standard errors are reported in parentheses below the parameter estimates. A Wald-test provides the χ^2 statistic for the null hypothesis while the *p*-value is given in square brackets below.

Table S.5. IV estimates of income underreporting relative to the comparison group. Households with two adults. Dependent variable: non-durable consumption

	(S3.1)	(S3.2)	(S3.3)	(S3.4)
	Full sample		**Matched sample**	
	Underreporting factor κ_g	**Underreporting factor s_g**	**Underreporting factor κ_g**	**Underreporting factor s_g**
Wage earners, share of business income > 0	1.278 (0.024)	0.218 (0.015)	1.662 (0.081)	0.398 (0.031)
Self-employed, share of business income > 0	1.518 (0.056)	0.341 (0.026)	1.506 (0.051)	0.336 (0.024)
Self-employed, share of business income = 0	1.114 (0.058)	0.103 (0.053)	1.100 (0.052)	0.091 (0.048)
Wald-test (H_0: $\kappa_1 = \kappa_2$)	21.38 [0.000]	..	2.76 [0.096]	..
Wald-test (H_0: $\kappa_3 = 1$)	3.92 [0.049]	..	3.71 [0.054]	..
No. of obs.	6016		3817	

Note: The comparison group of the estimations contains households who report that the household head is a wage earner and the household does not earn any business income. In the matched sample the group of wage earners with business income above 0 is matched to the group of self-employed with business income above 0 based on the share of business income. Bootstrapped standard errors are reported in parentheses below the parameter estimates. A Wald-test provides the χ^2 statistic for the null hypothesis while the *p*-value is given in square brackets below.

CURRICULUM VITAE

1. Personal data

Name	Merike Kukk
Date and place of birth	13.04.1976, Estonia
Citizenship	Estonia
Family status	Married, two children (born 2004 and 2008)
E-mail address	merike.kukk@ttu.ee

2. Education

Educational institution	Graduation year	Education (Field of study/degree)
Tallinn University of Technology	(2016)	Economics/ Doctor of Philosophy
Tallinn University of Technology	2004	Business administration/ Master of Arts
University of Tarty	1999	Economics/ Bachelor of Arts (*cum laude*)

3. Language competence/skills (fluent, average, basic skills)

Language	Level
Estonian	Native
English	Fluent
Swedish	Average
Russian	Average
German	Average
Spanish	Basic skills

4. Special courses

Period	Course
June 2015	Mastering 'Metrics: an Empirical Strategies Workshop, prof. Joshua D. Angrist (MIT) –University of Rome Tor Vergata
August 2014	Financial Markets and the Macro Economy, prof. Atif Mian (Princeton University), the Kiel Institute
June 2014	Panel Data Analysis, prof. Badi H. Baltagi (Syracuse University), University of Macedonia
July 2013	Discrete Choice Models for Cross Section and Panel Data, prof. William Greene (New York University),

	Economic Policies Research Unit, University of Minho
October 2012	Techniques and Applications of Robust Control, prof. Martin Ellison (University of Oxford), Danish Graduate Program of Economics, University of Copenhagen
July 2012	Applied Non-Parametric Methods, prof. Franco Peracchi (Tor Vergata University and EIEF), Barcelona Graduate School of Economics
May 2012	Introduction to Panel Data Econometrics, prof. Karl Taylor (University of Sheffield)
Sept-Oct 2010	Applied Time-series Econometrics, Juan Carlos Cuestas (University of Sheffield)

5. Professional employment

Period	Organization	Position
2016 –	Eesti Pank (national central bank)	Monetary policy and economic research department, researcher
2013, 2015	Eesti Pank	Visiting researcher
2012 –	Tallinn University of Technology	Research scientist
2005 – 2010	SEB	Deputy Head of Marketing and PR Division
2000 – 2005	SEB	Marketing Project Manager
1999 – 2000	Coca-Cola Baltic Beverages	Assistant/Product Manager

6. Research activity and thesis supervised

2014 – The best published article in social and human sciences in TUT (co-author K. Staehr)
2013 – Scientific award named after prof. Vello Vensel by Estonian Economic Association

Publications

Merike Kukk and Karsten Staehr (*forthcoming*): "Identification of Households Prone to Income Underreporting: Employment Status or Reported Business Income?", *Public Finance Review*.

Merike Kukk (*forthcoming*): "How Did Household Indebtedness Hamper Consumption during the Recession? Evidence from Micro Data", *Journal of Comparative Economics*.

Merike Kukk, Dmitry Kulikov and Karsten Staehr (*forthcoming*): "Consumption sensitivities in Estonia: Income shocks of different persistence", *Review of Income and Wealth*.

Merike Kukk and Karsten Staehr (2015): "Enhanced Fiscal Governance in the European Union: The Fiscal Compact", *Baltic Journal of European Studies*, Vol. 5, No.1, pp. 73-92.

Merike Kukk and Karsten Staehr (2014): "Income Underreporting by Households with Business Income. Evidence from Estonia", *Post-Communist Economies*, Vol. 26, No.2, pp. 257-276.

Madis Aben, Merike Kukk and Karsten Staehr (2012): "Housing Equity Withdrawal and Consumption Dynamics in Estonia 2002-2011", *Research in Economics and Business: Central and Eastern Europe*, Vol 4, No. 1, pp. 19-40.

Working Papers

Merike Kukk and Kasten Staehr (2015): "Macroeconomic Factors in Corporate and Household Saving. Evidence from Central and Eastern Europe", Working Paper series of Eesti Pank, No. 5/2015.

Merike Kukk (2015): "How Did Household Indebtedness Hamper Consumption during the Recession? Evidence from Micro Data", Italian Association for the Study of Economic Asymmetries, a/working paper No. 2015/05.

Merike Kukk (2014): "How Did Household Indebtedness Hamper Consumption during the Recession? Evidence from Micro Data", TUTECON Research Series, No. 5/2014.

Merike Kukk (2014): "Distinguishing the Components of Household Financial Wealth: the impact of Liabilities on Assets in European Countries", Working Paper series of Eesti Pank, No. 2/2014.

Merike Kukk and Karsten Staehr (2014): "Identification of Income Underreporting by the Self-Employed: Employment Status or Reported Business Income?", TUTECON Research Series, No. 1/2014.

Merike Kukk and Karsten Staehr (2013): "Income Underreporting by Households with Business Income. Evidence from Estonia", Working Paper series of Eesti Pank, No. 6/2013.

Merike Kukk, Dmitry Kulikov and Karsten Staehr (2012): "Consumption sensitivities in Estonia: Income shocks of different persistence", Working Paper series of Eesti Pank, No. 3/2012.

Other writings

Country Report for Estonia in "The Over-Indebtedness of European Households: Updated Mapping of the Situation, Nature and Causes, Effects and Initiatives for Alleviating its Impact", conducted for Directorate General Health and Consumers of the European Commission, December 2013.

Merike Kukk "How Much Do Expectations Matter in Estonia?" A Focus Feature in O. Blanchard and D. R. Johnson, textbook Macroeconomics, Global Edition 6/e, Pearson Education Inc., 2012.

Edmundz Rudzitis, Julita Varanauskiene, Triin Messimas, Merike Kukk. Baltic Household Outlook, biannual (April/Oct) analyses by SEB Group, since 2012.

Teaching

Economics II (Macroeconomics), Bachelor program in English
Macroeconomics, Master program in English
Foundations of fiscal policy, Bachelor program
Public Economics, Master program

7. Defended theses

The impact of marketing activities on the saving behaviour of households in Estonia, Master Thesis, supervisor Ann Vihalem, Tallinn University of Technology
The Currency Board and Its' Impact on the Economy, Bachelor Thesis, supervisor Janno Reiljan, University of Tartu

8. Main research topics

Household consumption, saving and borrowing behaviour, household tax behaviour

9. Additional information

2015 – Member of International Network for Economic Research
2013 – Member of European Economic Association
2012 – Member of Estonian Economic Association
2012 – Member of the work group of doctoral studies at TSEBA

ELULOOKIRJELDUS

1. Isikuandmed

Ees- ja perekonnanimi	Merike Kukk
Sünniaeg ja -koht	13.04.1976, Eesti
Kodakondsus	Eesti
Perekonnaseis	Abielus, kaks last (sünd. 2004 ja 2008)
E-posti aadress	merike.kukk@ttu.ee

2. Hariduskäik

Õppeasutus (nimetus lõpetamise ajal)	Lõpetamise aeg	Haridus (eriala/kraad)
Tallinna Tehnikaülikool	(2016)	Majandusteadus/ doktor
Tallinna Tehnikaülikool	2004	Ärikorraldus/ teadusmagister
Tartu Ülikool	1999	Majandusteadus/ bakalaureus (*cum laude*)

3. Keelteoskus (alg-, kesk- või kõrgtase)

Keel	Tase
Eesti keel	Emakeel
Inglise keel	Kõrgtase
Rootsi keel	Kesktase
Vene keel	Kesktase
Saksa keel	Kesktase
Hispaania keel	Algtase

4. Täiendusõpe

Õppimise aeg	Täiendusõppe korraldaja nimetus
Juuni 2015	Mastering ‘Metrics: kursus empiiriliste strateegiate kohta, prof. Joshua D. Angrist (MIT) – Rooma Tor Vergata Ülikool
August 2014	Finantsturud ja makroökonoomika, prof. Atif Mian (Princetoni Ülikool), Kieli Instituut
June 2014	Paneelandmete analüüs, prof. Badi H. Baltagi (Syracuse Ülikool), Makedoonia Ülikool
July 2013	Ristandmete ja paneelandmete diskreetse valiku mudelid, prof. William Greene (New York Ülikool), Minho Ülikool

October 2012	Usaldusväärsuse kontrolli võtted ja rakendused, prof. Martin Ellison (Oxfordi Ülikool), Kopenhaageni Ülikool
July 2012	Rakenduslikud mitte-parameetrilised meetodid. prof. Franco Peracchi (Tor Vergata Ülikool), Barcelona Majandusteaduse Doktorikool
May 2012	Sissejuhatus paneelandmete ökonomeetriasse, prof. Karl Taylor (Sheffieldi Ülikool)
Sept-Oct 2010	Rakenduslik aegridade ökonomeetria, Juan Carlos Cuestas (Sheffieldi Ülikool)

5. Teenistuskäik

Töötamise aeg	Tööandja nimetus	Ametikoht
2016 –	Eesti Pank	Rahapoliitika ja majandusuuringute osakonna uurija
2013, 2015	Eesti Pank	Külalisuurija
2012 –	Tallinna Tehnikaülikool	Teadur
2005 – 2010	SEB	Turundus- ja kommunikatsiooni-divisjoni direktori kt.
2000 – 2004	SEB	Turunduse projektijuht
1999 – 2000	Balti Coca-Cola Joogid	Assistent/ Tootejuht

6. Teadustegevus ja juhendatud lõputööd

2014 – TTÜ parim teadusartikkel sotsiaal- ja humanitaarteaduste valdkonnas (kaasautor K. Staehr)
2013 – prof. Vello Venseli nimeline Eesti Majandusteaduse Seltsi teauspreemia

Publikatsioonid

Merike Kukk and Karsten Staehr (ilmumas): "Sissetulekute alaraporteerimisele kalduvate majapidamiste identifitseerimine: kas tööalane staatus või raporteeritud ettevõtlustulu?", *Public Finance Review*.

Merike Kukk (ilmumas): "Kuidas pidurdas majanduslanguse ajal majapidamiste laenutase tarbimist? Tulemused mikroandmetest", *Journal of Comparative Economics*.

Merike Kukk, Dmitry Kulikov and Karsten Staehr (ilmumas): “Tarbimise tundlikkus erineva püsivusega sissetulekušokkidele kasutades raporteeritud sissetulekukomponente”, *Review of Income and Wealth*.

Merike Kukk and Karsten Staehr (2015): ”Fiskaalhaldus Euroopa Liidus: Fiskaalkokkulepe”, *Baltic Journal of European Studies*, Vol. 5, No.1, pp. 73 - 92.

Merike Kukk and Karsten Staehr (2014): “Sissetulekute osaline raporteerimata jätmine ettevõtlustuluga majapidamiste poolt Eestis”, *Post-Communist Economies*, Vol. 26, No.2, pp. 257-276.

Madis Aben, Merike Kukk and Karsten Staehr (2012): “Eluaseme väärtuse rahas väljavõtmine ja tarbimise dünaamika Eestis 2002-2011”, *Research in Economics and Business: Central and Eastern Europe*, Vol 4, No. 1, pp. 19-40.

Toimetised

Merike Kukk and Kasten Staehr (2015): “Makroökonoomilised faktorid ettevõtete ja majapidamiste säästmises Kesk-ja Ida-Euroopas”, Eesti Panga toimetised, nr. 5/2015.

Merike Kukk (2015): “Kuidas pidurdas majanduslanguse ajal majapidamiste laenutase tarbimist? Tulemused mikroandmetest”, Itaalia Majanduse Asümmeetriate Uurimise Selts, a/toimetised, nr. 2015/05.

Merike Kukk (2014): “Kuidas pidurdas majanduslanguse ajal majapidamiste laenutase tarbimist? Tulemused mikroandmetest”, TUTECON uurimistööde seeria, nr. 5/2014.

Merike Kukk (2014): “Vaadeldes majapidamiste finantsjõukuse komponente eraldi: laenude mõju finantsvaradele euroala riikides”, Eesti Panga toimetised, nr. 2/2014.

Merike Kukk and Karsten Staehr (2014): “Sissetulekute alaraporteerimise mõõtmine: tööturu staatus või raporteeritud ettevõtlustulu?”, TUTECON uurimistööde seeria, nr. 1/2014.

Merike Kukk and Karsten Staehr (2013): “Sissetulekute osaline raporteerimata jätmine ettevõtlustuluga majapidamiste poolt Eestis”, Eesti Panga toimetised, nr. 6/2013.

Merike Kukk, Dmitry Kulikov and Karsten Staehr (2012): “Tarbimise tundlikkus Eestis: erineva püsivusega sissetulekusokid”, Eesti Panga toimetised, nr. 3/2012.

Muud kirjutised

Ülevaade Eesti kohta raportis "Euroopa majapidamiste ülelaenamine: olukorra, põhjuste, tagajärgede ja kasutuselevõetud leevendusmeetmete kaardistamine", läbi viidud Euroopa Komisjoni Tervise ja Tarbijate Direktoraadi tellimusel, detsember 2013.

Merike Kukk "Kui olulised on Eestis majapidamiste ootused?" Fookusteema õpikus O. Blanchard ja D. R. Johnson, Makroökonoomika, rahvusvaheline trükk 6/e, Pearson Ed. Inc., 2012.

Edmundz Rudzitis, Julita Varanauskiene, Triin Messimas, Merike Kukk. „Ülevaade Balti majapidamiste kohta", kaks korda aastas (Aprill/Oktoober) analüüs SEB Grupi tellimusel, alates 2012.

Õppetöö

Majandusteooria II (Makroökonoomika), ingliskeelne bakalaureuseõpe
Makroökonoomika, ingliskeelne magistriõpe
Fiskaalpoliitika alused, bakalaureuseõpe
Avaliku sektori ökonoomika, magistriõpe

7. Kaitstud lõputööd

Eesti elanike säästukäitumise turunduslik mõjutamine, magistritöö, juh. Ann Vihalem
Valuutakomiteesüsteem ja selle mõju majandusele, bakalaureusetöö, juh. Janno Reiljan

8. Teadustöö põhisuunad

Majapidamiste tarbimis-, säästmis- ja laenukäitumine, majapidamiste maksukäitumine

9. Lisainfo

2015 – Majandusuuringute Rahvusvahelise Võrgustiku liige
2013 – Euroopa Majandusteadlaste Seltsi liige
2012 – Eesti Majandusteadlaste Seltsi liige
2012 – Majanduse doktoriõppe töörühma liige

DISSERTATIONS DEFENDED AT TALLINN UNIVERSITY OF TECHNOLOGY ON *ECONOMICS AND BUSINESS ADMINISTRATION*

1. **August Aarma**. Segmented Analysis of Bank Customers and Banking Information: Estonian Case. 2001.
2. **Enn Listra**. The Development and Structure of Banking Sector: Retail Banking in Estonia. 2001.
3. **Tatyana Põlajeva**. The Comparative Analysis of Market's Attractiveness. 2001.
4. **Tuuli Tammeraid**. Modeling Flow of Funds for Estonia. 2002.
5. **Ivo Karilaid**. The Choice in General Method for Investment and Performance Evaluation. 2002.
6. **Hele Hammer**. Strategic Investment Decisions: Evidence from Survey and Field Research in Estonia. 2003.
7. **Viljar Jaamu**. The Methods and Instruments for Solving the Banking Crisis and Development of the Banking Sector in Estonia. 2003.
8. **Katri Kerem**. From Adoption to Relationships: Internet Banking in Estonia. 2003.
9. **Ly Kirikal**. Productivity, the Malmquist Index and the Empirical Study of Banks in Estonia. 2005.
10. **Jaanus Raim**. The PPP Deviations between Estonia and Non-Transitional Countries. 2006.
11. **Jochen Sebastian Heubischl**. European Network Governance – Corporate Network Systematic in Germany, the United Kingdom and France: an Empirical Investigation. 2006.
12. **Enno Lend**. Transpordiühenduse ja logistikasüsteemi interaktsioon (Saaremaa ja Hiiumaa näitel). 2007.
13. **Ivar Soone**. Interrelations between Retail Service Satisfaction and Customer Loyalty: A Holistic Perspective. 2007.
14. **Aaro Hazak**. Capital Structure and Dividend Decisions under Distributed Profit Taxation. 2008.
15. **Laivi Laidroo**. Public Announcements' Relevance, Quality and Determinants on Tallinn, Riga, and Vilnius Stock Exchanges. 2008.
16. **Martti Randveer**. Monetary Policy Transmission Channels, Flexibility of the Economy and Future Prospects of the Estonian Monetary System. 2009.

17. **Kaire Põder**. Structural Solutions to Social Traps: Formal and Informal Institutions. 2010.
18. **Tõnn Talpsepp**. Investor Behavior and Volatility Asymmetry. 2010.
19. **Tarmo Kadak**. Creation of a Supportive Model for Designing and Improving the Performance Management System of an Organisation. 2011.
20. **Jüri Kleesmaa**. Economic Instruments as Tools for Environmental Regulation of Electricity Production in Estonia. 2011.
21. **Oliver Parts**. The Effects of Cosmopolitanism on Estonian and Slovenian Consumer Choice Behavior of Foreign *versus* Domestic Products. 2011.
22. **Mart Nutt**. Eesti parlamendi pädevuse kujunemine ja rakendamine välissuhetes. 2011.
23. **Igor Novikov**. Credit Risk Determinants in the Banking Sectors of the Baltic States. 2011.
24. **Mike Franz Wahl**. Kapitaliühingute lõppomanike alusväärtuste ja tahte uurimine ning omanikkonna tüpoloogia konstrueerimine. 2011.
25. **Tobias Wiebelt**. Impact of Lease Capitalization on the Development of Accounting Standards: A Theoretical Research. 2012.
26. **Sirje Pädam**. Economic Perspectives on Environmental Policies: The Costs and Benefits of Environmental Regulation in Estonia. 2012.
27. **Juhan Värk**. Venemaa positiivse hõlvamise poliitika ja teiste välispoliitiliste liinide mõjud Eesti-Vene suhetele aastail 1991–2011. 2012.
28. **Mari Avarmaa**. Implications of Capital Structure and Credit Constraints for Company Performance: A Comparative Study of Local and Multinational Companies in the Baltics. 2012.
29. **Fabio Filipozzi**. The Efficiency of Interest Rate and Foreign Exchange Markets in the Euro Area and Central and Eastern Europe. 2012.
30. **Aleksei Netšunajev**. Developments and Determinants of Intra-Industry Trade in the Baltic States. 2012.
31. **Aleksandr Miina**. Critical Success Factors of Lean Thinking Implementation in Estonian Manufacturing Companies. 2012.
32. **Angelika Kallakmaa-Kapsta**. Before and After the Boom: Changes in the Estonian Housing Market. 2013.
33. **Karen Voolaid**. Measurement of Organizational Learning of Business Schools. 2013.
34. **Archil Chochia**. Models of European Integration: Georgia's Economic and Political Transition. 2013.

35. **Hannes Ling**. Developing an Assessment Measure for Enhancing Entrepreneurship Education through a Metacognitive Approach. 2013.
36. **Marina Järvis**. Assessment of the Contribution of Safety Knowledge to Sustainable Safety Management Systems in Estonian SMEs. 2013.
37. **Aleksandr Michelson**. Destination Branding and Constructing Symbolic Capital in the Reproduction of Heritage Space: the Case of UNESCO Designated Hanseatic Old Towns. 2014.
38. **Alar Kolk**. Co-evolution of Capabilities' and Alliance Portfolios: Multinational Firms in Global ICT Industry. 2014.
39. **Svetlana Raudonen**. Impact of Corporative Taxation on Foreign Direct Investments: Evidence from the European Union. 2014.
40. **Riina Koris**. Customer Orientation at a Higher Educational Institution: The Perspective of Undergraduate Business Students in Estonia. 2014.
41. **Merle Rihma**. Ethics Audit: A Management Tool for Assessing of Corporate Social Responsibility and Preventing Ethical Risks. 2014.
42. **Anu Virovere**. The Role of Management Values, Knowledge Management and Conflict Management for Improvement of Organisational Sustainability. 2015.
43. **Kristina Hunke**. Conceptualisation and Management of Green Transport Corridors. 2015.
44. **Eneken Titov**. Management Paradigm Values in Real and Propagated Level as Prerequisites of Organisational Success. 2015.
45. **Siiri Same**. Conceptualization of Experience Marketing and Country Branding from a Marketing Management Perspective. 2015.
46. **Helbe Põdder**. Combining Work and Family Life – A Comparative Perspective on Changes in Parental Roles in Estonia, Finland and Sweden. 2015.
47. **Katrin Toompuu**. The Full Costing Model and Its Implementation at Universities: The Case of Tallinn University of Technology. 2015.
48. **Õnnela Paas**. Development of the Safety Management System at Enterprises. 2015.
49. **Iivi Riivits-Arkonsuo**. Consumer's Journey as Ambassador of Brand Experiences. 2015.
50. **Liis Talpas**. The Theoretical Bases for the Preparation of Financial Statements for SMEs: the Case of Estonia. 2016.
51. **Olga Nežerenko**. Formal Cluster Formation in the Development of the Transport Sector in the Baltic Sea Macro-Region. 2016.