A New Method to Estimate Tax Evasion Using Financial Institution Lending: The Case of Greece

Nikolaos Artavanis
Virginia Polytechnic Institute and State University

Adair Morse
University of Chicago

Margarita Tsoutsoura
University of Chicago

October 19, 2011

Abstract

We introduce a new methodology to estimate tax evasion which is based on households' ability to borrow. In many developed economies, a large percentage of the population earns at least some income in actions unreported to tax authorities. Pervasive tax evasion implies that banks adapt credit models, adjusting households' reported income to reflect their perception of true income. We use detailed individual loan application data in Greece to quantify the extent of tax evasion country-wide. Our estimates suggest 18 to 26 billion euros in taxable income goes unreported and the foregone revenues for the tax authorities amount to 27% to 41% of the deficit. We then document the incidence of tax evasion by income class and occupation. We find that the upper income accounts for a large portion of tax evasion implied by bank adaptation. Doctors and many professional service occupations are the largest offenders. Finally, we discuss the occupational distribution in light of paper trail mandates and theories as to the optimal level of tax subsidies.

1 Introduction

Most developed countries have large informal (or semi-formal) economies across a wide spectrum of professions. The average size of the shadow economy across all developed countries is estimated to be more than 12% of GDP (Sneider and Enste, 2003). For Mediterranean countries, estimates are much larger, perhaps due to a cultural norm of informality. The cost to informality is the loss of taxable revenues for the government. In some cases, as in the case of Greece, the tax base is insufficient to cover government obligations, as the recent sovereign debt crisis in Europe demonstrates.

Estimating tax noncompliance is a challenging task, since tax evasion is an act that is meant to be hidden. We introduce a new methodology based on households' ability to borrow. Our innovation uses the fact that financial institutions "adapt" to informality by providing formal lending off informal income.[1] When individuals apply for loans, banks observe reported income (that which is filed to tax authorities) and then use models to translate reported income to an ability-to-pay-equivalent to true income based on what they know about tax evasion for each occupation. We develop a new methodogy to estimate tax evasion based on observed lending. We use individual-level loan application data from one of the ten large banks in Greece and we quantify the extent of tax evasion for the whole country of Greece. We then apportion tax noncompliance across occupations and income class.

Our contribution builds on the second of the two main microeconomic methodologies to estimate tax evasion. The first strand of the literature uses "direct" methods, employing audits of tax returns (i.e. studies that use the IRS's Taxpayer Compliance Measurement Program data). The second set of methods indirectly estimates evaded income from observed expenditure data. Two prominent studies in this strand of literature are Pissarides and Weber (1989) and Feldman and Slemrod (2007). Pissarides and Weber develop the consumption-based approach and use food expenditure survey data and reported income in UK to estimate the underreporting of British self-employed. They find that on average the true income of self-employed is 1.55 times their reported income. Feldman and Slemrod (2007) use the relationship between reported char-

[1] We borrow the term "adapt" from Harberger (2006) who discusses customs tax evasion and institutional adaptation.

itable contributions and reported income, and find tax evasion among self-employed, nonfarm small-business and farm income are 1.54, 4.54 and 3.87 times, respectively, reported income. The consumption-based methodology has been applied in a host of settings. Andreoni et al. (1998) and Slemrod and Yitzaki (2002) offer a comprehensive review of the literature.[2]

In some sense, by focusing on borrowing, our method is the flip-side of the consumption approach in that we measure tax evasion off the mechanism of intertemporal smoothing of consumption; namely, debt. This is really not a a fair characterization of our methodology. If we were just estimating off those with negative net worth, a fair criticism would have been that not all individuals are at the lifecycle stage of borrowing to smooth consumption. However the use of debt products is not at all limited to negative net worth individuals. First, credit cards are used pervasively in the economy as a method of exchange. In Greece, the rise of credit card use has been relatively recent (the last decade), which is a convenient fact for us since we employ application data from this time period. In addition, it is well documented that household borrow and save at the same time, primarily through mortgages. Thus, the use of our very large sample of individuals using a variety of debt products (including zero-borrowing debt products), combined with detailed, zip-code level statistics from the tax authority allows us to use debt to speak to the economy at large.

Using debt and household level data to infer true income has some benefits. First, we are able to estimate tax evasion even for wage workers. More fundamentally, we are able to exploit the fact that tax evasion is endemic in certain occupations. Banks have formal models to inflate reported income to their best estimate of true income *by occupation and employment status* (self employed or not). We know from conversations with multiple banks in Greece as well as other banks in South Europe that bank lending adaptation by occupation is pervasive. In uncollateralized lending, such as the bulk of consumer lending, banks are only concerned with individual cash flows, in particular income, being sufficient to repay loan servicing consistently.

[2] A separate literature relies on macroeconomic approaches to estimate the size of the black economy.The most common approaches are consumption methods (e.g., as in the electricity approach of Lacko (1999)) and the currency demand approach (Cagan (1958), Tanzi (1983)). These methods are best suited to estimate the size of the informal economy, which emcompass but are not specific to (income) tax evasion. Sneider (2002) gives an overview of these methods, discussing their benefits and limitations and higlighting differences between the black economy estimates and income tax evasion.

If anything, banks make conservative estimates by occupation-employment status group. The fact that adaptation happens implies that tax evasion is an intergrated aspect in the occupation-self employment status groups. In other words, rather than estimating tax evasion through the heterogeneity of individual consumption and then applying it to occupations to observe the mean tax evasion by occupation, we use a method that relies on occupational standards. If occupations have standards of tax evasion, then it is both theoretically and in practice easier to analyze optimality of the extent to tax evasion by type of occupation-employment class and, importantly, solutions for governments like the one in Greece of today.

The individual-level lending data we use to identify tax evasion is household lending – credit cards, terms loans, mortgages, overdraft and the like. From application data, we have rich information on reported income, total debt outstanding, occupation, employment status (self- employed or wage earner), credit history and demographics for many tens of thousands of applicants. We know the zip code of the borrowers, which allows us to control for (often hereditary or family) wealth effects which might be picked up in consumption studies. We will be able to control for fixed effects that absorb risks of different occupations.

Our methodolgy replicates a ratio or a scoring model for a finance provider to determine the credit capacity of a household borrower using all of these variables. Of course, the underlying identification problem for the debt capacity model (and the goal of the paper) is that income observed is reported income, not true income. Our main approach uses the structure of the financial institutions adaptation model. The model generates unbiased occupation multiplier estimates for the self-employed (specific to each occupation) under the assumption that wage earners are not tax evading. To the extent that wage earners do tax evade, which is known to be true in certain occupations, our estimates are downward biased. We implement a more general version of this model to allow for wage earner tax evasion using a structural model with reported income, consumption and debt measures of the latent variable of true income and life cycle predictors as structurally related to income.[3]

Our bank model results find 18 to 27 billion euros in evaded taxable income just for the self-employed. The tax base in Greece for 2010 is 86 billion euros; thus our magnitude is very meaningful. At the tax rate of 40%, the foregone tax revenues would account for 27% to 41%

[3] These results do not appear in this draft.

of the budget deficit shortfall in 2010.

At the occupation level, we find a higher tax evasion multiple for the self-employed dentists and veterinarians, educators, professional sevices and professional admininstrators, lawyers, doctors, and engineers. In terms of euros, the largest tax evaders among the self- employed are doctors, dentists and veterinarians, professional services and lawyers. We reconcile this cross section of offenders with a legislative bill that targeted nine select occupations (doctors, dentists, veterinarians, lawyers, architects, engineers, topographer engineers, economists, firm consultants and accountants) and with (forthcoming) notions of paper trail (capturing the ease of evading) and with theories of where tax evasion might be optimal.

We also can break our estimates of tax evasion by income. We find that the top decile of reported income earners account for 25-50 percent of tax evaded income.

Our study concludes with two (forthcoming) simulation exercises to offer remedies. Documenting that certain occupations tax evade does not solve the problem of collection. The approach of tax authorities actually using observable bank data would would never work, even if such data were allowed to be observable to tax authorities. Individuals would shift borrowing patterns. However, our approach allows us to suggest two solutions using our estimates. First, tax authorities could impose minimum taxes at the occupation-zip code level, with the burden of proof for those falling below the limit transfered to the individual. We simulate how much additional tax revenues Greek authorities could collect under this *presumed income* model, following the examples of other European countries, notably Italy. Second, we propose an occupation establishment tax (as implemented by city authorities world-wide), discuss applicability to specific occupations and simulate revenue using this method.

The remainer of the paper is as follows. Section 2 introduces our rich bank and tax authority data, and provides summary statistics by self-employed or not and by occupation. Section 3 discusses prior methods. Section 4 lays out our methodologies, and in particular, the care with which we must approach removing inconsistencies in estimating the sensitive of debt to reported (versus true) income. Section 5 reports results and interprets magnitudes at the economy-level. Section 6 is our policy simulation. Section 7 concludes.

2 Data

We use a proprietary panel dataset from a large Greek commercial bank. The bank has tens of thousands lending accounts, and its branches cover the entire country. There are 10 major banks in Greece that cover at least eighty percent of the market. The bank that provides the data is one of these banks. The dataset contains the universe of consumer loan applications that the bank received, both approved and rejected. Furthermore for the approved applications we have detailed monthly information on the performance of each individual loan. The consumer loan applications (term loans, open loans, credit cards, overdrafts) cover the period January 2003 to February 2011, while the mortgage applications cover the period January 2006 to February 2011. The applications have detailed information both about the loan requested and the applicant. We have information on the date of the application, the branch office and the loan officer who evaluated the application, the purpose of the loan, the requested and approved loan amount and loan duration, and whether the loan was approved or rejected, as well as the reason for the rejection. Moreover we have detailed data on every person who was involved in the loan, as applicant, co-applicant or guarantor. We observe the person's demographic characteristics, homeownership, existing prior relationship with the bank, income as well as debt outstanding with other lenders. Furthermore, we have very detailed information about the applicant's occupation and employment type (wage worker, self-employed, pensioner etc.). Occupationis verified by the bank since the involved parties ned to provide employer verification letter. Self-employment is defined by our datasource using the main source of income and is verified by the tax return.

There are a number of unique characteristics of these data that make them suitable for the purpose of this study. The applications have data on the personal and family income of the applicants. The income is verified by the bank using the tax returns of the applicants. The data are comprehensive and highly reliable as they comply with the Basel II requirements.

We supplement these data with detailed zipcode level data from the Greek tax authorities. For every zipcode we have deciles of income for all tax filers as well as their classification in four employment categories : Merchants and Small Business Owners, Agriculture, Wage Earners and Self-Employed. We supplement the rich dataset from the bank with the detailed income deciles per zipcode data from the tax authorities, to weight our sample and make it representative to the population. For our analysis , we exclude students, pensioners and unemployed, since our goal is to focus on the active workforce.

Tables 1a and 1b summarize by employment status the main variables used in the analysis. From Tables 1a and 1b we observe that around that around forty percent of the sample has prior relationship with the bank, while the average length of this relationship is four years. Table 2a and 2b show reported income and total debt outstanding for self-employed and wage earners in the consumer loans and mortgage sample respectively. The total debt outstanding is the total debt with our bank as well as with other lenders. In our analysis we use total debt outstanding as a measure of debt capacity. In reality debt capacity and debt outstanding are different and furthermore this difference may be correlated to employment status, as self-employed might be more likely to max out their debt capacity (in the methodology section, we ignore this difference, just referring to debt capacity). Later in the analysis, we handle this potential measurement error of the dependent variable empirically by isolating situations in which the borrower is likely to be maxed out on credit limit.Tables 2a and 2b show a stricking difference in the different debt outstanding between wage workers and self-employed. This difference is consistent almost across all occupations. So even in a naive comparison of average income and debt outstanding, the data show that self-employed have much higher levels of debt, although they do not have higher reported incomes. Of course we are not able to derive conclusions from such a naive comparison, since, among other reasons, the distributions of income and debt outstanding might be different for self-employed and wage workers, self-employed may have different risk profile and growth prospects and there are also differences among professions. In the next section we describe our empirical methodology that would address these challenges.

3 Methodology

Our goal is to estimate the extent to tax evasion and financial institution adaptation by occupation and employment status. We use a variety of methods to ensure robustness and to offer lower bounds to the estimates.

3.1 Lending Adaptation

We implement a framework to document tax evasion starting from the debt extended by the bank, a fully observable variable, starting from bank models of credity capacity. In other words, we use model of how much debt a financial institution is willing to extend to individuals to infer true income from observed debt. With true income in hand, calculating tax evasion. Bank models can be simple ratio models or more complex scoring models. We generate estimates based on both. First, it is helpful to lay out some notation.

We denote debt capacity of individual i by d_{ijk}. An important dimension of our analyses is going to be employment; thus, we subscript occupation and employment status (private sector wage earner or self-employed) with j and k respectively. Debt capacity is a function of true income ξ_{ijk}, as well as other covariates such as occupation, credit history and demographics. Because of tax evasion, the bank only observes reported income, y_{ijk}, which is necessarily not greater than true income since nobody intentionally overreports income to the tax authorities. Financial institution adaptation is the procedure by which the bank adjusts lending to reflect true income.

3.1.1 Ratio Rule Model

A standard practice is for lenders to use ratio thresholds to decide whether a loan application should be accepted or rejected. The ratio on which we focus is debt-to-income. Our application data provide an individual's overall debt outstanding, including debt obligations to other lenders. Our acceptance/rejection data have a variable that codes the reason for the rejections. We use the bank rejections of loan application under the rejection categories "Relationship of loan to income" and "Relationship of payments to income" as situations when debt capacity is maxed out. Our measure of the debt-to-income rule, *ratio*, is the ex ante debt outstanding relative to income for individuals seeking a loan who were rejected for one of these two reason codes.

$$ratio_j = \underset{i \in jk_{wage}}{median}\left(\frac{d_{ijk_{wage}}}{y_{ijk_{wage}}}\right) \text{ for the individuals } i \text{ rejected for a consumer loan.}$$

A few comments are, first, that we calculate the ratio only for wage earners. Following the literature (Pissarides and Weber, 1989; Feldman and Slemrod, 2007), this assumes that wage earners do not tax evade. (Some of our subsequent analyses relax this assumption; the social norm in Greece is for wage earners to earn informal wages in after-work hours in some occupations.) Second, although it is commonly assumed that the debt-to-income rule is fixed across individuals, we know from discussions with the bank that a more accurate depiction of the bank's implementation of ratios allows them to vary by occupation and age to control for varying default risk. We calculate the rejection ratio as the median outstanding debt-to-income ratio by age bracket-occupation combination. We do this along a split of pre-crisis and post-crisis (split on October 31, 2009), to incorporate adjustments made by the bank after the liquidity crisis.

Finally, we also assume in this calculation is that the ex ante debt outstanding prior to asking for the loan is the full debt capacity. We know that the debt outstanding plus the debt requested is larger than the debt capacity. The true debt capacity is somewhere in between, but the ex ante debt level is our best guess without imposing subjectivity. Because this in between range is large for mortgage applications, we limit this analysis to consumer credit applications, where the new loan amounts requested are very small relative to the debt outstanding.

With $ratio_j$ in hand, it is straightforward to apply the rule that debt cannot exceed the ratio for all other individuals, namely, the self-employed. Apply the statistics to all self-employed individuals in occupation j leads to estimates of true income as

$$\widehat{\xi}_{ijk_{self}} = d_{ijk_{self}} \cdot \frac{1}{ratio_j}.$$

The final step is to subtract out reported income from estimated true income and taking the mean over occupation j self-employed individuals.

It is certainly possible and perhaps even likely that self-employed individuals have different thresholds of borrowing relative to income. However, the estimates provided by the ratio model are conservative as we explain below. First when we apply the ratio model to self-employed we assume that people have maxed out their debt. If in reality people are more conservative and do not max out their debt capacity, then our estimates underestimate the extend of tax evasion.[4] Furthermore the use of the ratio model assumes that the rejection debt-income ratio

[4] Due to the conservative nature of the method, for people that have low levels of debt relative to their debt

fot self-employed is similar to the rejection ratio for selfemployed. In the data we observe that the rejection debt-income ratio for self-employed is higher across professions, which means that our estimates are conservative.

Nevertheless, we move to a more formal credit scoring model of debt capacity.

$$\widehat{Y}_{ij,k\in self} = dc_{ij,k\in self} \cdot \frac{1}{ratio_j}.$$

3.2 Debt Capacity Scoring Model

Our formal debt capacity model replicates the scoring process of lenders to determine credit limits. Throughout our scoring analysis, we utilize two debt capacity models. We begin with the *simple debt capacity model* that most closely approximates our understanding of the bank's adaptation. We then add more econometric rigor to the model and present the *occupation effects debt capacity model.*

The underlying debt capacity equation for the simple debt capacity model is:

$$d_{ijk} = \beta_o \xi_{ijk} + X_i \beta_1 + \beta_2 I_i^{SE} + \varepsilon_{ijk} \tag{1}$$

Following the prior notation, the dependent variable d_{ijk} is individual i's debt capacity level measured in euros, with j indexing occupation, and k indexing self-employment versus wage worker. The most important component of a debt capacity model is income. The model includes ξ_{ijk}, true income, which is often different from observed income y_{ijk}. Independent variables in X are other factors used by the bank to determine debt capacity, namely,the individual's relationship with the bank, credit history, household demographics and stability factors. Although we later look at time dimensions to events in Greece, we write the model as a cross section. We embed time fixed effects in X to incorporate supply changes to the credit model, particularly after the liquidity crisis.

capacity, the estimated income might be lower than their declared income. Since in reality is very unlikely that people declare to the tax authorities higher income than their real income we restrict tax evasion to be nonnegative. We present results both unrestricted and restricted to forcing tax evasion to be nonnegative.

Whether or not the individual is self-employed also matters for determinations of credit, even after incorporating true income into the model. From the bank's perspective, repayment risk might be higher for the self-employed than wage workers. However,the bank would balance repayment risk against prospects for profits from additional services which could be provided to the self-employed if income were to growth. It is unclear which effect would dominate. We incorporate a self employment effect, $\beta_2 I_i^{SE}$, to absorb the overall implication to being self-employed.

If we knew true income, ξ_{ijk}, could we write the debt capacity scoring model structurally, without the error term ε_{ijk}? The answer is not quite. We observe all the data in the applications and past performance of loans, and thus we have all the hard information the bank uses. However, we have no reason to believe that the bank process is not free from soft information entering loan decision process at the local level.[5] On average, the soft information utilization should be just noise across the large number of loan officers, given that individual branches follow the company-level guidelines. The other unknown is the exact form of the scoring model and which variables are actually utilized. We know that the debt capacity model is linear in functional form. Thus, we use a "kitchen sink" approach in including the variables nonparametrically (using buckets of levels rather than as continuous variables) where appropriate and letting the estimation load on which however many variables are important.

One clear omission in the principle of including all possible variables in this *simple debt capacity model* is that we have not allowed for an occupation effect. Income may not be a sufficient statistic for occupation-specific risks incorporated by the bank. In addition, self-employment may matter in a way specific to each occupation. For example, self-employment in construction or retail occupations may involve added risk relative to doctor self-employment income, since the former depends more on market conditions relative to private medical practices.Thus we write the *occupation effects debt capacity model* as:

$$d_{ijk} = \beta_o \xi_{ijk} + X_i \beta_1 + \mu_{.j.} + \mu_{..k} + \mu_{.jk} + \varepsilon_{ijk}. \quad (2)$$

We denote an employment status effect by $\mu_{..k}$, occupation fixed effects by $\mu_{.j.}$ and fixed effects

[5] In a future draft, we hope to remove loan officer effects to absorb any systematic patterns in soft information use. However, we doubt this is a necessary step given the very large number of loan officers in the sample and the centralized system of hard information used by the bank.

for self-employment specific to the occupation by $\mu_{\cdot jk}$. The *occupation effects debt capacity model* is a blatently better econometric model than the *simple debt capacity model*. The reason we use the latter at all is that we believe the bank implementation of income adaptation, below, follows the simple model. Nevertheless, we run all results through both models.

The problem with equation (1) is that income ξ_{ijk} is not generally observable and using reported income, y_{ijk}, in place of true income would create a bias in the estimate of β_0. Furthermore, the divergence of reported and true income is almost surely correlated with other estimators, which would create havoc in the entire estimation if we do not handle the .

The banks are not naive and instead have a process to adapt credit scoring to reflect their best guess of true income. The bank discussed with us with the essence of their credit scoring methodology (but not sufficiently so to use their scores or know their parameters). From this and discussions with other banks, we are able to infer that the primary variation of adaptation is along dimensions of employment; namely, occupation and employment status. In particular, another bank told us point-blank that they use an occupation multiplier to scale-up reported income to their estimate of true income $m_{jk}(\cdot)\, y_{ijk}$. The adaptation model $m_{jk}(\cdot)$ is a simple multiplier on reported income:

$$\xi_{ijk} = m_{jk}(\cdot)\, y_{ijk} + \nu_{ijk}, \tag{3}$$

with a series of occupation multipliers, specific to self employment (I^{SE}) or not ($1 - I^{SE}$):

$$m_{jk} = I^{SE}\lambda_j + \left(1 - I^{SE}\right)\lambda_j \phi_j^{wage}. \tag{4}$$

In equation (4), λ_j is the self employed income multiplier and $\lambda_j \phi_j^{wage}$ is the wage earner income multiplier.

It is useful to think about the residual term ν_{ijk}. This residual is not an error in the implementation of adaptation. The bank has standard rules that dictate such adjustments down to the branches. Rather, the residual reflects the fact that the adaption model is the best guess as to true income on average for the occupation-employment status group. Other individual characteristics surely affect the propensity to tax evade. For example, the extent to tax evasion may have locational or age components. What is convenient for us is that even if the bank is wrong on average with its adaption model, inserting the adapted true income transformation $\xi_{ijk} = \left[I^{SE}\lambda_j + \left(1 - I^{SE}\right)\lambda_j \phi_j^{wage}\right] y_{ijk}$ remove the bias in the debt capactiy equation, since debt capacity is determined by the bank. To this end, if the bank is wrong on average, our ultimate analysis of $\widehat{\xi}_{ijk}$ and tax evastion $\widehat{\xi_{ijk} - y_{ijk}}$ can be interpreted as true income implied by the bank and tax evasion implied by the bank. However, since the bank's incentives with regards to income are, if anything, conservative (penalites for overestimating income are greater than rewards for underestimating income), our analysis is also conservative. Given the size and experience of our bank, our prior is that the bank probably has a pretty good sense of the underlying income of the population, and we procedure using such terminology.

To implement , we begin our analysis with the standard assumption that $m_{j,wage} = 1$; i.e., for wage earners, reported and true incomes are equivalent.[6] In such a setting, our estimating equations for the two models reduce to:

Simple Debt Capacity Model:

$$d_{ijk} = \beta_o^S \left(1 - I_i^{SE}\right) y_{ij,wage} + \left(\beta_o \lambda_j\right)^S I_i^{SE} y_{ij,SE} + X_i \beta_1^S + \beta_2^S I_i^{SE} + \zeta_{ijk}^S. \tag{5}$$

Occupation Effects Debt Capacity Model:

$$d_{ijk} = \beta_o^O \left(1 - I_i^{SE}\right) y_{ij,wage} + \left(\beta_o \lambda_j\right)^O I_i^{SE} y_{ij,SE} + X_i \beta_1^O + \mu_{\cdot\cdot k} + \mu_{\cdot j \cdot} + \mu_{\cdot jk} + \zeta_{ijk}^O. \tag{6}$$

We have added S and O superscripts to differentiate the model parameters. Under the assumptions that we have not left out essential factors in the adaption model, the β estimates will be consistent: $\zeta_{ijk}^m = \beta_0^m \nu_{ijk} + \varepsilon_{ijk}^m$ for $m \in \{S, O\}$ is uncorrelated with any of the other variables. If occupation effects matter in the debt capacity scoring, beyond adaptation, then the *simple model* could be inconsistent, and the *occupation effects debt model* is more efficient. However, the interpretation of the adaptation parameter is more intuitive in the *simple debt capacity model.*

[6] We will relax this assumption in a future draft.

In the *simple debt capacity model*, the estimates $\left\{\widehat{\beta_o^S}, \left\{\widehat{\beta_o^S\lambda_{j=1}^S}, ..., \widehat{\beta_o^S\lambda_{j=M}^S}\right\}\right\}$ allow us to identify $\left\{\lambda_{j=1}^S, ..., \lambda_{j=M}^S\right\}$ across the M occupations. These λ's are the direct multiplier by which the bank adjusts reported income by occupation for the self employed. Thus, tax evasion for occupation j is calculated as $\frac{1}{N_i}\sum_{i=1}^{N_i}\left[\left(\lambda_j^S - 1\right) y_{ijk_{SE}}\right]$.

In the *occupation effects debt model*, we have to take into account the fixed effects for occupation and occupation-employment status. In this model, the estimator β_o captures the effect of true income over and above the occupation-employment status mean effect on debt capacity.

4 Results

4.1 Results from the Ratio Model of Debt Capacity

Table 3 presents the results of the ratio model. As we described earlier, this method is based on the rejection debt-to-income ratios that banks implement in their lending decisions. Details on the method are provided in the methodology section. We apply the calculated ratios to the self-employed to back out their estimated income based on their debt outstanding. The first column of Table 3 presents the average declared income of self-employed based on their tax returns.Due to the conservative nature of the method, for people that have low levels of debt relative to their debt capacity, the estimated income might be lower than their declared income. Since in reality is very unlikely that people declare to the tax authorities higher income than their real income we restrict tax evasion to be nonnegative (in columns 2 and 3). In Columns 4 and 5 we present the unrestricted results.Our results show pervasive underreporting of income across professions. In terms of euros, the highest tax evasion is observed among factory[7] owners, pharmacists and doctors. In terms of the extent of underreporting, retail and business services on average appear to declare only 54% of their incomes, while dentists and veterinarians and

[7] The results of the ratio model should be interpreted with caution for those professions that the nature of the job of the selfemployed is very different from that of the wage worker. In our main model our specification allows to overcome these challenges.

construction report less than sixty percent of their incomes. As we analyze in the methodology the estimates provided by the ratio model are conservative.

4.2 Credit Scoring Model Results

Table 4 presents our first results from the credit scoring model under the assumption that wage earners do not tax evade. Table 4 focuses on consumer loan estimation. The dependent variable is total debt outstanding for each individual. Credit history variables include whether the pulling credit report was authorized, the source of the loan, whether collateral was pledged, and the length of history with bank. We include risk variables of type of loan, length of time in job and length of time in residence. Demographic controls are spouse occupation, the sum of spouse and guarantor income, indicators for marital status, the number of dependents, and ten-year age brackets. To isolate supply effect, we include time dummies. All estimations are weighted to the population using the tax authority data described in the data section.

Columns 1-3 report the coefficients of interest from the *occupation effects debt scoring model.* In particular, column 1 presents the estimates on the reported income of the self employed by occupation. Column 2 presents the occupation fixed effects. And column 3 presents the fixed effects for occupation crossed with self employment. In column 4, we report estimates from the *simple debt capacity score model.*

The first order of business is to check whether the estimates in the more econometrically robust model (columns 1-3) can be replicated in the simpler model whose interpretation has a closer appeal to our understanding of what the bank does. Comparing coefficients in columsn 1 and 4 reveal a high degree of similarities, with the exception of Agriculture. Thus, we feel comfortable with focusing our interpretation on column 4.

Panel B aids the interpretation, converting the across-the-board significnat coefficients into the λ multipliers and implied tax evasion. The multiplier is large (around 2.5-3.0X reported income) for Dentists & Veterinarians, Educators, Financial Services, Professional Services, Accountants, Medical Other, Lawyers and Doctors. In euro terms, tax evasion is the most for Doctors, Pharmacists, Professional Services, Dentists and Lawyers. The overall tax evaded income in Greece implied by the population weighted lambda is 18.8 billion euros.

We mentioned in the data section that debt capacity and debt outstanding are empirically not the same. To address this, we limit our sample to situations in which they are likely to be sufficiently close as to eliminate any systematic biases.

Columns 5 and 6 of Table 4 Panel A present results for only those loan applications for restructuring or overdraft loans. The goal of these columns is to ascertain whether the magnitude, significance and ranking of occupation tax evasion multipliers remains the same when focusing on a group of individual with maxed-out debt capacities. Our estimates in columns 5 and 6 suggest just this.The largest offenders are Accountants & Notaries, Doctors, Educators, Professional Services and Lawyers. The effect for Dentist and Pharmacists do not appear to be robust across specifications. New to the list of the worst offenders are Factory Owners, Retail and Small Business.

Table 5 builds on principle of the last few columns of Table 4 panel A.We use the sample of individuals with approved mortgages as people who have likely taken their debt capacity to the limit. Financial institutions make mortgage decisions with loan-to-value models. Thus, the dependent variable for the Table 5 estimation is aprroved mortgage plus other outstanding debt relative to the value of the house. We put other debt in numerator since in Greece lenders can break up mortgages (like in the United States) into separate loans.[8] Using an approved mortgage sample is nice, in addition to providing a situation in which households are likely to be constrained, in that we are able to implement a different model (a collateral-based model) to see how robust our results are to a different model.

Table 5, panel A shows the mortgage results for the occupation effects and simple credit capacity models. Although our sample of approved mortgages is less than five percent of the sample of consumer loan applications, we have a sufficiently large sample to identify both models, albeit with less significant results. The largest tax evading occupations are financial services, brokers, factory owners, professional administrators and doctors. In money terms (panel B), doctors again are largest offenders, with the exception of financial services which brings in a huge amount of tax evaded euros according to estimates.

The overall economic magnitude is in line with sample of all consumer loans. Our mortgage estimates suggest that 26.8 billion euros of taxable income goes unreported.

[8]Note that we are careful to not double count loan applications due to mortgage dispersion.

5 Policy Implication: A Simulation Exercise

Our results suggest that from 18 to 27 billion euros of taxable income go unreported. With a tax rate of 40% in Greece, this suggests an additional 8-10 billion euros of tax revenue which could be collected. However, collecting said revenue is a non-trivial task. In Greece, as is elsewhere in Europe (list of countries), the tax authorities cannot get access to individual banking records to implement a tax collection based on our direct strategy of adjusting taxable income for each individual based on debt. However, our results can be applied directly using a *presumed tax* method. Such a model is implemented for certain sectors in Italy and elsewhere.[More Forthcoming. We think it is important that this actually is an effective policy elsewhere.].

The idea is that we know the tax evasion by occupation for each zip code. In Greece, there are approximately 1,500 zip codes for a population of 5.5 million working households. (The zip codes a very precise.) We can simulate how much tax the government could collect if it were to mandate a minimum tax reporting by occupation-zip code using statistics from our estimation by occupation. Such a policy might induce migration, but presumably, proximate zip codes would have a similar *presumed tax* schedule, and the schedules can be geographically smoothed to ensure minimal border incentives.

6 Income Distribution of Tax Evasion

The common understanding of tax evasion is that it is an upper income phenomenon. We can study the incidence of tax evasion by income decile. Figures 1 and 2 present these results. As the figures show, it is indeed true that the rich have hide larger fraction of their income from the tax authority. However, particularly in the constrained model, the middle class are very active tax evaders as well, with average magnitudes of tax evading self employed being in the twenties of thousands of euros per person.

7 Conclusion

Using individual-level household lending data, we develop a new methodology to estimate tax evasion based on the household's ability to borrow. Our methodology is based on the fact that due to the pervasive informality financial institutions need to adapt their lending and infer true income based on the reported income.

Using our methodology we are able to quantify the extend of tax noncompliance overall in Greece. Our estimates suggest that the unreported taxable income is 5 percent to 9 percent of GDP. Furthermore using our method we analyze how the tax noncompliance varies by occupation and income class. We find that the upper income accounts for a large portion of tax evasion implied by bank adaptation. Doctors and many professional service occupations that are largely service oriented and have less paper trail mandates are the largest offenders.

8 References

Andreoni, James; Erard, Brian and Feinstein, Jonathan. "Tax Compliance." Journal of Economic Literature, 1998, XXXVI(June 1998), pp. 818-60.

Cagan, Phillip. "The Demand for Currency Relative to the Total Money Supply." Journal of Political Economy, 1958, 66(4), pp. 303-28.

Chetty, Raj and Saez, Emmanuel. "Optimal Taxation and Social Insurance with Endogenous Private Insurance." American Economic Journal: Economic Policy, 2010, 2(2), pp. 85-114.

Feldman, Naomi E. and Slemrod, Joel. "Estimating Tax Noncompliance with Evidence from Unaudited Tax Returns." Economic Journal, 2007, 117, pp. 327-52.

Harberger, Arnold. "Reflections on Corruption: An Economist's Perspective," Strategic Objective or Activity: 905-101/Economic Growth. U.S. Agency for International Development, 2006.

Lacko, Maria. "Hidden Economy - an Unknown Quantity? Comparative Analysis of Hidden Economies in Transition Countries, 1989-95." Economics of Transition, 2000, 8(1), pp. 117-49.

Pissarides, Christopher A. and Weber, Guglielmo. "An Expenditure-Based Estimate of Britian's Black Economy." Journal of Public Economics, 1989, 39(1), pp. 17-32.

Schneider, Friedrich and Enste, Dominik H. "Shadow Economies: Size, Causes and Consequences." Journal of Economic Literature, 2000, 38(1), pp. 77-114.

Slemrod, Joel and Yitzhaki, Shlomo. "The Costs of Taxation and the Marginal Efficiency Cost of Funds." IMF Staff Papers, 1996, 43(1), pp. 172-98.

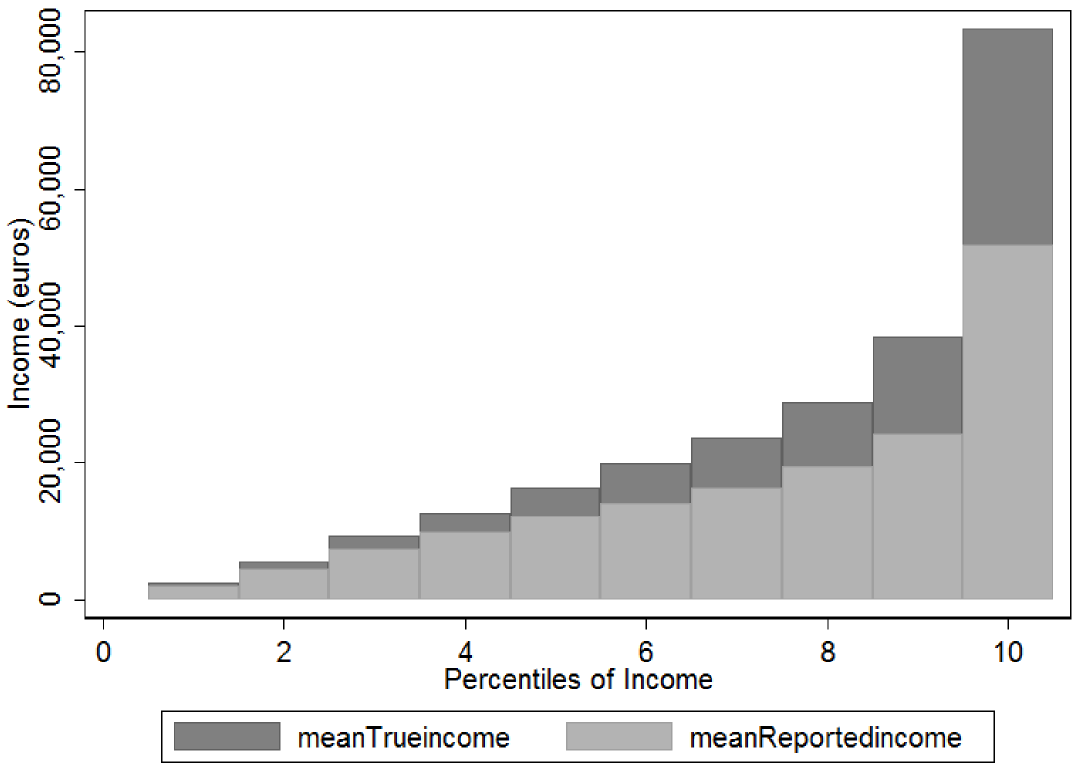

Figure 1: Incidence of Tax Evasion by Income Decile.
Figure 1 plots the tax evasion estimates from the simple debt capacity model with the full sample of loan applications (Table 4, column 4 estimates).

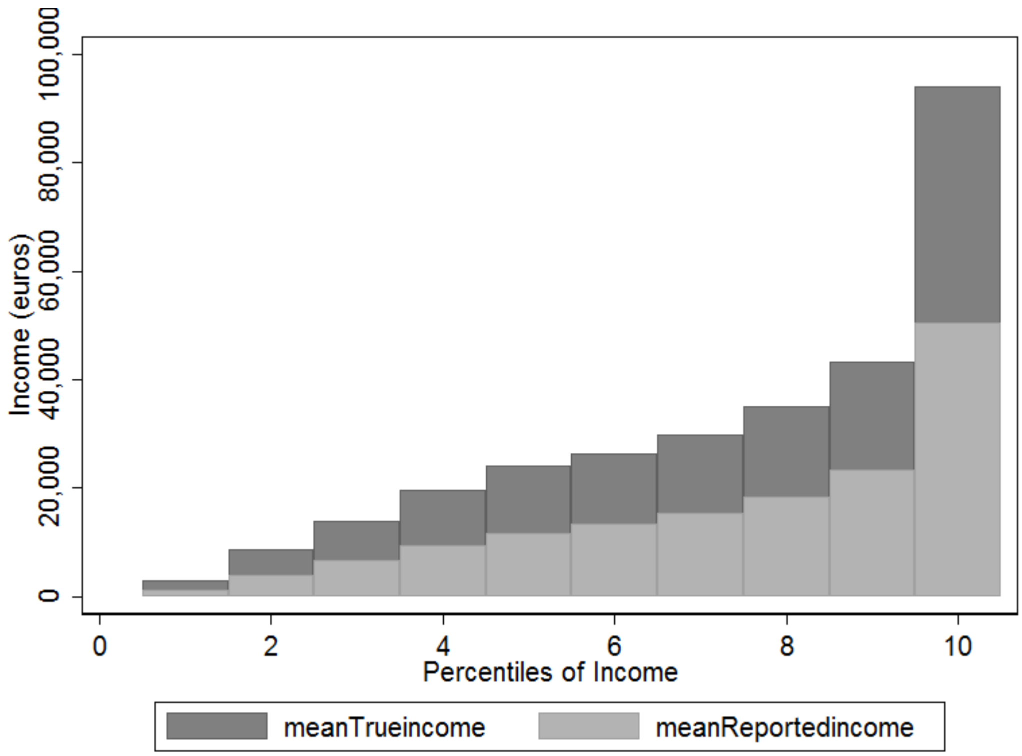

Figure 2: Incidence of Tax Evasion by Income Decile, constrained consumer sample.
Figure 2 plots the tax evasion estimates from the simple debt capacity model with the sample of loan applications who are taking either overdraft or restructuring loans (Table 4, column 6 estimates).

Table 1a. Income, Debt Outstanding and Demographic Characteristics by Employment Status

This table presents average income and debt outstanding as well as demographic characteristics for wage workers and self-employed in the consumer loans sample. Personal income and spouse income are verified using tax returns. Prior debt outstanding is the debt outstanding at the time of the application and includes debt obligations to other lenders. The loan approved amount and the percantage of requested loan approved are calculated only for approved applications. The sample is weighted to the population using the tax authority data described in the data section.

Consumer Loans Sample	**Wage Workers**	**Self employed**	**Total**
Personal Income (€)	15,049	16,398	15,303
Spouse Income (€)	6,932	5,945	6,721
Loan Requested Amount (€)	6,346	7,866	6,632
Loan Approved Amount (€)	8,277	10,506	8,620
Percentage of Loans Approved	0.52	0.45	0.50
Age	41.69	45.03	42.32
Years in Address	12.99	15.32	13.42
Years in Job	8.45	11.17	8.96
Marital Status	0.60	0.70	0.62
Depedent Persons	0.44	0.53	0.46
Existing prior relationship with the bank	0.45	0.43	0.45
Years of cooperation with bank	3.51	3.98	3.60

Table 1b. Income, Debt Outstanding and Demographic Characteristics by Employment Status

This table presents average income and debt outstanding as well as demographic characteristics for wage workers and self-employed in the mortgage sample. Personal income and spouse income are verified using tax returns. Prior debt outstanding is the debt outstanding at the time of the application and includes debt obligations to other lenders. The summary statistics are calculated only for approved loans, since the loan-to-value variable is meaningful only for approved loans. The sample is weighted to the population using the tax authority data described in the data section. In the mortgage sample there is no information about the years of cooperation with the bank.

Mortgages Sample	**Wage Workers**	**Self employed**	**Total**
Personal Income (€)	21,411	13,389	20,631
Spouse Income (€)	8,505	5,763	8,239
Loan Requested Amount (€)	75,842	91,137	77,328
Loan Approved Amount (€)	99,387	114,804	100,885
Percentage of Loans Approved	0.83	0.73	0.81
Total Loans-to-Commercial Value	0.95	1.15	0.97
Age	44.01	46.84	44.29
Years in Address	10.99	14.03	11.28
Years in Job	10.75	13.30	11.00
Marital Status	0.94	0.95	0.94
Depedent Persons	0.74	0.73	0.74

Table 2a. Income and Debt Outstanding by Occupation and Employment Status/Consumer Loans

This table presents average personal income and debt outstanding by occupation and employment status in the consumer loans sample. Personal income and spouse income are verified using tax returns. Debt outstanding includes debt obligations to other lenders. All estimations are weighted to the population using the tax authority data described in the data section.

	Income (€)			Debt Outstanding (€)		
Occupation	**Wage Workers**	**Self-employed**	**Total**	**Wage Workers**	**Self-employed**	**Total**
Accountants & Notaries	17,619	20,026	18,449	26,077	40,369	31,007
Agriculture	10,556	11,444	10,996	7,059	12,047	9,533
Artists & Athletes	16,026	16,024	16,025	19,144	25,265	21,473
Brokers	18,134	21,184	20,076	23,166	34,784	30,562
Construction	11,435	16,971	12,336	8,489	20,771	10,488
Dentists & Veterinarian	21,356	17,509	18,869	32,816	38,847	36,715
Doctor	27,684	32,957	28,780	38,937	52,631	41,783
Educator	19,520	17,191	19,350	27,322	36,109	27,963
Engineering & Building Professional	19,734	20,097	19,891	19,634	33,455	25,610
Factory Owners and Workers	15,220	17,815	15,610	16,314	30,626	18,462
Financial Services	21,942	20,837	21,364	36,292	41,854	39,200
Laborer	12,337	12,511	12,363	8,814	10,926	9,124
Lawyers	28,787	20,253	23,387	45,980	39,580	41,930
Medical Other	16,238	17,836	16,330	22,813	30,261	23,243
Others	14,117	15,841	14,411	12,494	23,856	14,434
Pharmacist	21,314	42,928	31,728	23,120	63,463	42,560
Professional Admin & Others	20,449	19,839	20,318	26,984	29,663	27,560
Professional Services	21,439	19,993	20,847	28,045	35,023	30,903
Retail & Hotel	12,575	19,702	15,262	16,890	31,283	22,316
Scientist	18,151	16,626	17,739	20,394	25,527	21,781
Small Business Others	14,862	15,591	14,971	17,249	19,123	17,528
Transport	23,687	18,413	23,419	25,176	21,827	25,005
Total	15,050	16,398	15,303	16,501	23,434	17,806

Table 2b. Income and Debt-to-Commercial Value by Occupation and Employment Status/Mortgages

This table presents average personal income and total debt to commercial value of real estate property at the time of application by occupation and employment status in the mortgage sample. Personal income and spouse income are verified using tax returns. Total debt outstanding includes debt obligations to other lenders. All estimations are weighted to the population using the tax authority data described in the data section.

	Income (€)			Total Loans-to-Commercial Value		
Occupation	**Wage Workers**	**Self-employed**	**Total**	**Wage Workers**	**Self-employed**	**Total**
Accountants & Notaries	21,881	16,169	20,543	0.96	1.19	1.01
Agriculture	12,454	11,576	12,035	0.87	1.00	0.93
Artists & Athletes	27,275	13,328	25,016	0.89	1.08	0.92
Brokers	28,526	13,486	22,942	1.09	1.33	1.18
Construction	13,650	10,836	13,281	0.82	1.07	0.85
Dentists & Veterinarian	31,016	13,562	22,377	1.05	1.04	1.04
Doctor	34,113	29,361	33,526	1.04	1.24	1.07
Educator	22,512	8,807	22,105	0.95	1.16	0.95
Engineering & Building Professional	28,068	19,267	25,594	0.94	1.14	0.99
Factory Owners and Workers	19,422	13,171	18,941	0.94	1.16	0.95
Financial Services	27,281	11,550	22,394	1.22	1.65	1.35
Laborer	17,393	9,174	17,183	0.89	0.85	0.89
Lawyers	39,329	17,227	30,723	1.13	1.01	1.09
Medical Other	18,935	13,515	18,786	1.01	1.12	1.02
Others	19,324	11,898	15,399	0.88	1.22	1.06
Pharmacist	32,045	14,912	25,275	1.15	1.12	1.14
Professional Admin & Others	27,165	15,443	25,940	0.98	1.15	1.00
Professional Services	35,410	21,909	33,196	0.99	1.18	1.02
Retail & Hotel	15,668	11,023	14,690	0.90	1.16	0.95
Scientist	23,538	19,467	23,153	0.92	1.31	0.95
Small Business Others	21,359	13,502	21,045	0.94	1.15	0.95
Transport	34,833	11,298	34,070	1.00	0.98	1.00
Total	21,411	13,389	20,631	0.95	1.15	0.97

Table 3. Ratio Model of Debt Capacity - Estimation of Tax Evasion of Self-Employed by Occupation

The table demonstrates the estimated mean income, estimated mean tax evasion and estimated percentage of tax evaders for self-employed individuals in each occupation. The estimates have been obtained using the ratio model of debt capacity. Following the literature (Pissarides and Weber, 1989; Feldman and Slemrod, 2007), we assume that wage earners do not tax evade. The ratio method cannot be applied to observations with zero debt oustanding and therefore they are excluded. Due to the conservative nature of the method, for people that have low levels of debt relative to their debt capacity, the estimated income might be lower than their declared income. Since in reality is very unlikely that people declare to the tax authorities higher income than their real income we restrict tax evasion to be nonnegative (in columns 2 and 3). In Columns 4 and 5 we present the unrestricted results. All estimations are weighted to the population using the tax authority data described in the data section.

Occupation	Declared Personal Income (€)	Estimated Income (€)	Estimated Tax Evasion (€)	Estimated Income / Unconstrained (€)	Estimated Tax Evasion / Unconstrained (€)	Estimated Tax Evaders (%)
	1	2	3	4	5	6
Accountants & Notaries	20,863	43,060	22,197	35,205	14,342	0.39
Agriculture	12,151	26,376	14,225	21,856	9,705	0.40
Artists & Athletes	16,681	36,171	19,496	30,072	13,397	0.39
Brokers	22,354	51,034	28,680	43,390	21,037	0.42
Construction	18,941	43,590	24,649	36,537	17,596	0.42
Dentists & Veterinarian	18,133	41,717	23,625	35,230	17,138	0.39
Doctor	35,033	65,653	30,619	50,309	15,276	0.35
Educator	18,159	39,362	21,203	33,079	14,920	0.40
Engineering & Building Professional	21,740	52,582	30,842	44,266	22,526	0.40
Factory Owners and Workers	18,839	184,830	165,991	178,117	159,278	0.43
Financial Services	22,088	55,632	33,544	49,019	26,931	0.47
Laborer	14,033	31,373	17,341	26,303	12,270	0.42
Lawyers	21,157	36,765	15,544	28,002	6,782	0.35
Medical Other	19,037	38,182	19,157	30,819	11,795	0.38
Others	16,923	39,915	22,991	34,133	17,210	0.43
Pharmacist	46,310	116,012	69,687	96,944	50,619	0.39
Professional Admin & Others	20,860	51,946	31,086	43,002	22,142	0.33
Professional Services	21,064	47,062	26,039	39,577	18,554	0.41
Retail & Hotel	20,839	49,946	29,108	42,205	21,367	0.43
Scientist	17,714	49,585	31,842	42,118	24,375	0.32
Small Business Others	16,780	33,606	16,826	27,015	10,235	0.40
Transport	20,663	37,053	16,390	27,587	6,925	0.38

Table 4: Debt Capacity Models

Panel A: Estimation. Description appears on Panel B page.

Sample:	All Consumers				Constrained Borrowers	
Model:	Occupation Effects (A single estimation columns 1-3)			Simple Debt Capacity	Occupation Effects	Simple Debt Capacity
	1	2	3	4	5	6
Personal Income (b_0)	0.316***			0.353***	0.190***	0.213***
	[0.00755]			[0.00796]	[0.0215]	[0.0227]
Self Employment (b_2)	5,904***			416.4*	33,764**	6,908***
	[490.1]			[213.0]	[15,609]	[1,914]
Variable with occupation index:	$y_{ijk} I_i^{SE}$	μ_j	μ_{jk}	$y_{ijk} I_i^{SE}$	$y_{ijk} I_i^{SE}$	$y_{ijk} I_i^{SE}$
Accountants & Notaries	0.758***	omitted	-7,372***	0.921***	0.949*	1.415***
	[0.105]		[2,249]	[0.0653]	[0.518]	[0.298]
Agriculture	0.314***	-16,341***	-3,763***	-0.00251	0.283**	-0.125*
	[0.0342]	[528.7]	[620.6]	[0.0153]	[0.120]	[0.0742]
Artists & Athletes	0.596***	-3,925***	-9,323***	0.601***	0.898**	0.849**
	[0.128]	[797.8]	[2,222]	[0.0746]	[0.433]	[0.341]
Brokers	0.517***	-1,360	-6,631***	0.625***	1.432***	1.347***
	[0.0810]	[1,128]	[2,086]	[0.0574]	[0.281]	[0.229]
Construction	0.463***	-14,221***	-1,713**	0.412***	0.770***	0.717***
	[0.0363]	[526.0]	[747.5]	[0.0257]	[0.147]	[0.126]
Dentists & Veterinarian	1.104***	1,719	-14,383***	1.127***	0.387	0.948
	[0.309]	[1,699]	[5,581]	[0.118]	[0.501]	[0.673]
Doctor	0.737***	7,243***	-10,940***	0.889***	0.924**	1.081***
	[0.0860]	[938.1]	[2,947]	[0.0578]	[0.413]	[0.238]
Educator	1.008***	-3,292***	-8,972***	1.045***	0.436	1.300***
	[0.206]	[564.0]	[3,469]	[0.0992]	[0.531]	[0.369]
Engineering	0.732***	-5,099***	-5,741***	0.784***	0.711***	0.996***
	[0.0638]	[643.3]	[1,433]	[0.0435]	[0.236]	[0.170]
Factory Owners & Workers	0.627***	-8,970***	-4,098***	0.637***	1.466***	1.372***
	[0.0541]	[533.5]	[1,089]	[0.0335]	[0.263]	[0.195]
Financial Services	0.890***	6,522***	-15,067***	1.025***	0.503	0.840***
	[0.0873]	[1,154]	[2,227]	[0.0520]	[0.312]	[0.211]
Laborers		-12,532***				
		[539.0]				
Lawyers	0.788***	11,273***	-19,832***	0.903***	1.110***	0.729***
	[0.107]	[1,716]	[2,737]	[0.0672]	[0.224]	[0.172]
Medical Other	0.955***	-4,823***	-9,331***	0.927***	1.366	1.320**
	[0.137]	[576.4]	[2,301]	[0.0864]	[0.904]	[0.576]
Others	0.515***	-10,192***	-3,735***	0.501***	0.592***	0.614***
	[0.0287]	[521.2]	[667.6]	[0.0188]	[0.173]	[0.132]
Pharmacist	0.426***	-4,266**	14,799***	0.752***	-0.0261	1.597***
	[0.105]	[1,763]	[5,230]	[0.0638]	[0.670]	[0.289]
Professional Admin & Others	0.529***	-2,573***	-7,511***	0.604***	1.400**	1.067***
	[0.0797]	[694.3]	[1,748]	[0.0556]	[0.596]	[0.371]
Professional Services	0.896***	4,359***	-14,458***	0.985***	1.799***	1.675***
	[0.147]	[1,208]	[3,090]	[0.0889]	[0.658]	[0.398]
Retail & Hotel	0.392***	-6,729***	-574.5	0.515***	1.281***	1.376***
	[0.0214]	[540.8]	[648.9]	[0.0184]	[0.444]	[0.272]
Scientist	0.447***	-4,981***	-5,159	0.583***	0.486	-0.417
	[0.148]	[1,101]	[3,291]	[0.115]	[1.421]	[0.453]
Small Business Others	0.498***	-6,133***	-8,501***	0.457***	0.971***	0.920***
	[0.0222]	[520.2]	[590.4]	[0.0157]	[0.226]	[0.158]
Transport	0.425***	-2,707***	-13,686***	0.377***	-0.0312	0.0486
	[0.0970]	[591.4]	[1,731]	[0.0751]	[0.124]	[0.0980]
R-squared	0.171			0.178	0.276	0.271

Table 4, Panel B

Panel A presents the results from the credit scoring models in which we assume that wage earners do not tax evade. Table 4 focuses on consumer loan estimation. The dependent variable is total debt outstanding for each individual. Credit history control variables include whether the pulling credit report was authorized, the source of the loan, whether collateral was pledged, and the length of history with bank. We include risk variables of type of loan, length of time in job and length of time in residence. Demographic controls are spouse occupation, the sum of spouse and guarantor income, indicators for marital status, the number of dependents, and ten-year age brackets. To isolate supply effect, we include time dummies. All estimations are weighted to the population using the tax authority data described in the data section. Columns 1-3 report the coefficients of interest from the occupation effects debt scoring model. In particular, column 1 presents the estimates on the reported income of the self employed by occupation. Column 2 presents the occupation fixed effects. And column 3 presents the fixed effects for occupation crossed with self employment.

Panel B presents the estimate from Table 4, Panel A, Column 4 (the simple debt capacity model) for the estimatted lambda multiplier, the implied tax evasion and true income by profession. At the bottom of the table, the lambdas are averaged with population weights and applied ot overall Greece statistics from the tax authority to arrive at an estimate of the loss in taxable income implied by the estimates.

	Lambda	**Tax Evasion**	**Reported Income**	**True Income**
Accountants & Notaries	2.61	39,426	24,559	63,986
Agriculture	-0.01	-11,522	11,441	-81
Artists & Athletes	1.70	13,031	18,627	31,658
Brokers	1.77	20,136	26,232	46,368
Construction	1.17	2,744	16,619	19,363
Dentists & Veterinarian	3.19	43,352	19,803	63,155
Doctor	2.51	63,975	42,230	106,205
Educator	2.96	39,351	20,101	59,453
Engineering & Building Professional	2.22	31,273	25,671	56,944
Factory Owners & Workers	1.80	16,280	20,315	36,595
Financial Services	2.90	48,689	25,614	74,303
Laborer	n/a	n/a	13,347	n/a
Lawyers	2.56	38,866	24,974	63,840
Medical Other	2.62	35,085	21,621	56,706
Others	1.42	7,763	18,600	26,363
Pharmacist	2.13	47,532	42,155	89,688
Professional Admin & Others	1.71	16,450	23,184	39,634
Professional Services	2.79	44,390	24,847	69,237
Retail & Hotel	1.46	8,899	19,456	28,355
Scientist	1.65	12,198	18,769	30,968
Small Business Others	1.29	5,186	17,764	22,950
Transport	1.07	1,233	18,617	19,850

Population weighted lambda:	1.51	
Percentage of Population Self-Employed	0.42	
Taxable Income for Self-Employed 2010	35.7	billion euros
Tax Evasion Economy-Wide Estimate 2010	18.2	billion euros

Table 5: Debt Capacity Models in Mortgage Sample

Panel A: Estimation. Description appears on Panel B page.

Sample:	**Mortgage Customers**	
Model:	**Occupation Effects**	**Simple Debt Capacity**
	1	**2**
Personal Income (b_0)	1.934***	2.117***
	[0.278]	[0.274]
Self Employment (b_2)	-0.06474	0.1177***
	[0.138]	[0.0123]
Variable with occupation index:		
Accountants & Notaries	-0.919	0.589
	[1.972]	[1.736]
Agriculture	13.19***	-0.820
	[3.237]	[2.110]
Artists & Athletes	0.0422	-0.612
	[4.023]	[3.154]
Brokers	3.706	8.331*
	[4.929]	[4.709]
Construction	5.002***	4.170***
	[1.583]	[1.337]
Dentists & Veterinarian	4.289	1.564
	[3.425]	[2.628]
Doctor	6.598***	5.811***
	[1.920]	[1.506]
Educator	1.449	3.654
	[5.697]	[5.225]
Engineering	4.350**	4.375***
	[1.827]	[1.506]
Factory Owners & Workers	7.413***	6.892***
	[1.743]	[1.628]
Financial Services	19.68**	26.12***
	[8.029]	[5.963]
Lawyers	2.550	0.508
	[2.057]	[1.702]
Medical Other	4.321	3.657
	[5.958]	[5.071]
Others	3.769*	5.108**
	[1.952]	[2.005]
Pharmacist	-1.540	1.829
	[2.704]	[2.658]
Professional Admin & Others	8.364	6.653*
	[5.273]	[3.975]
Professional Services	0.604	2.728
	[2.418]	[2.149]
Retail & Hotel	2.290**	2.630***
	[1.059]	[1.010]
Scientist	3.791	3.359
	[6.909]	[6.282]
Small Business Others	2.480***	2.607***
	[0.724]	[0.684]
Transport	4.688**	3.470
	[2.235]	[2.120]
R-squared	0.073	0.069

Table 5, Panel B

Panel B presents the estimate from Table 4, Panel A, Column 4 (the simple debt capacity model) for the estimatted lambda multiplier, the implied tax evasion and true income by profession. At the bottom of the table, the lambdas are averaged with population weights and applied ot overall Greece statistics from the tax authority to arrive at an estimate of the loss in taxable income implied by the estimates.

	Lambda	Tax Evasion	Reported Income	True Income
Accountants & Notaries	0.278	not significant		
Agriculture	-0.387	not significant		
Artists & Athletes	-0.289	not significant		
Brokers	3.935	46,624	15,883	62,507
Construction	1.970	12,461	12,846	25,307
Dentists & Veterinarian	0.739	not significant		
Doctor	2.745	60,618	34,732	95,350
Educator	1.726	not significant		
Engineering & Building Professional	2.067	24,039	22,531	46,570
Fabrication	3.256	35,653	15,806	51,458
Financial Services	12.340	163,130	14,385	177,515
Lawyers	0.240	not significant		
Medical Other	1.727	not significant		
Others	2.413	19,599	13,869	33,468
Pharmacist	0.864	not significant		
Professional Admin & Others	3.143	38,378	17,911	56,289
Professional Services	1.289	not significant		
Retail & Hotel	1.242	3,145	12,980	16,125
Scientist	1.587	not significant		
Small Business Others	1.232	3,742	16,160	19,901
Transport	1.639	not significant		

Population weighted lambda:	1.75	
Percentage of Population Self Employed	0.42	
Taxable Income for Self Employed 2010	35.7	billion euros
Tax Evasion Economy-Wide Estimate 2010	26.8	billion euros