Abstract

I estimate a life-cycle labor supply model to quantify individuals’ exposure to permanent earnings risk. I find that permanent earnings risk in the US has been on the rise since the early 2000s and has taken a marked hike during the financial crisis of 2008. In contrast, the insurance effect of the progressive tax and transfer system, which mitigates this risk, has remained flat. I estimate the progressivity of the tax and transfer system using a power function approximation. This progressivity parameter is sufficient to identify the insurance effect of the tax and transfer system. Further, when progressivity is shut down, the model features only 5% less insurance. Earnings risk could have been reduced to pre-crisis levels by increasing progressivity substantially, lowering the progressivity parameter from the observed level of 0.93 to 0.78.
1 Introduction

Over their life-cycle individuals experience changes in the remuneration of their work. Some of these changes in earnings are expected, some are not. The unforeseen changes constitute risk that alters forward-looking working, saving, and consumption decisions. The class of models investigating these responses typically assume an incomplete financial market and frame saving with a non-contingent bond as the central device to insure against income risk. In this class of models the consumption reaction to a permanent change in income should be strong, even one-to-one if the utility function does not exhibit positive prudence.\footnote{See \cite{Krueger2007} for an illustration with a specific income process and quadratic utility. Further, see \cite{Attanasio2010} for a comprehensive survey on the issue of consumption responses to changes in the income process. Precautionary saving, which operates when prudence is positive, can smooth the consumption profile.} However, as for example \cite{Blundell2008} document, pass-through of a permanent shock from income to consumption is not one-for-one empirically, which is generally referred to as \textit{excess smoothness}. This phenomenon has lead researchers to explore other sources of insurance that mitigate the pass-through, such as precautionary saving, labor supply, labor supply of the secondary earner, and progressive taxation (\cite{Blundell2015,Blundell2016a,Heathcote2014}).

The current paper sets out to answer two major questions related to the riskiness of earnings:

1. How has permanent earnings risk developed over the last decade in the US and to what extent has progressive taxation insured against this risk?

2. What is the insurance effect of the current and alternative tax schedules?

The analysis builds on a life-cycle model of labor supply, which features risk originating in hourly wages. The model is tractable and allows for the estimation of the relevant parameters giving the evolution of risk and the behavioral parameters of labor supply in two steps: 1. Estimation of growth in log wages and log hours, derived from the first order condition for labor supply, which yields residual wage and residual hours growth. 2. Method of moments estimation using the second moments of residual wage and residual hours growth, which yields the relevant parameters to measure risk and pass-through to earnings. This approach has successfully been employed by both \cite{Blundell2016a} and \cite{Heathcote2014} for the purpose of exploring and comparing different insurance channels. The approach relies on the approximation of the labor supply equation and the life-time budget constraint.\footnote{The accuracy of these approximation methods has been explored in \cite{Domeij2006,Blundell2008,Blundell2016a} among others. I will not examine the properties of these approximations.} Income taxation, like in previous contributions (\cite{Blundell2016a} and \cite{Heathcote2014}) is modeled by means of a power function approximation. Two novelties of the current paper are the integration of tax deductions and that I find a form of the approximation that is compatible with nonlabor income. In both cases the forms are chosen to maintain the tractability of the first order condition approach. On the empirical side, to quantify risk and insurance, I follow the standard approach of estimating shock variances and calculating their pass-through to earnings.
use a novel approach to assess how changes in the tax code alter the insurance provided by progressive taxation. First, I slightly modify the tax code of the tax calculator and compute the new distribution of incomes. Second, using this new distribution of incomes, I estimate the parameters of the tax function approximation and calculate the relative change to the base level.

I employ US data from the Panel Study of Income Dynamics (PSID) from 1998 to 2014, which means I can cover the influence of the US financial crisis of 2008 on my measure of risk. The PSID has been used extensively in other studies of this issue and enables me to embed my findings in the literature.

As concerns the research questions, I find a moderate rise in the permanent component of wage risk until 2006 (18% increase in the standard deviation) and a strong increase from 2006 to 2008 (14% increase in the standard deviation) that persists until 2010 and only partially reverts in 2012. In studying the pass-through of this “gross” risk onto earnings I find that progressive taxation played a minor mitigating role. On average about a 5% decrease in pass-through compared to the case with no progressive taxation. I cannot detect a major change in the estimated progressivity and, in turn, insurance around the 2008 crisis. The rise in gross risk after the crisis transfers to net earnings. Finally, I examine a counterfactual calculation, of how high the top tax rate needed to be to return earnings risk to pre-crisis levels. I conclude that an increase of about 56% of the actual rate would have achieved this goal.

As concerns the modeling of the tax function, I find that tax deductions are of minor importance in shaping how progressive the tax system is. However, they do have an influence on the parameters of the approximated function. Further, when investigating the approximation of net income, I find that the estimates are sensitive to whether one fits the relationship in levels or logs. The approximation in levels provides a better fit and implies less progressivity. Overall, I find that the fitted power function performs relatively well. I calculate a fiscal gap per tax unit, the difference between actual and fitted tax burden, of roughly -650$ implying that the model slightly underpredicts liabilities.

In the next section I provide an overview of the related literature. In section 3 I relate the model set up and then focus on the approximation of the retention function in sections 4 to 6.2. I introduce the PSID data in section 5. Section 7 presents the results of estimating the life-cycle model and section 8 contextualizes these results in terms of the pass-through of wage risk to earnings risk and conducts the policy exercise regarding the top tax rate.

2 Taxation as Insurance

The idea that taxation can act as an insurance mechanism against risky income flows dates back to the 1980s with foundational contributions by Varian (1980) and Eaton and Rosen (1980). While Varian (1980) considers a dynamic model, in which an individual may self-insure against income risk via saving, Eaton and Rosen (1980) explicitly model labor supply with uncertainty in wages but neglect dynamics. Both Eaton and Rosen (1980) and Varian (1980) come to the conclusion that taxation can be desirable if an individual faces shocks and is risk averse. Varian (1980) considers an optimal nonlinear tax schedule
with the finding that a more progressive tax schedule is optimal, when the Arrow-Pratt measure of absolute risk aversion is increasing. Under the imposition of a utility function that features constant relative risk aversion (CRRA) marginal tax rates increase in income, given that the coefficient of relative risk aversion is larger than 1.

Low and Maldoom (2004) make the connection between the two papers and examine optimal taxation when labor income is risky and both labor supply and savings are choice variables of the individual. The fundamental trade-off is between the incentive effect on labor supply stemming from income uncertainty and the benefit of social insurance that derives from lowering the variance of net income. They determine that the trade-off is parametrized by the ratio of prudence to risk aversion.

Several recent empirical studies like Blundell et al. (2016a), Blundell et al. (2015), Heathcote et al. (2014), and Heathcote et al. (2017a) seek to estimate the degree of insurance stemming from sources like savings, (family) labor supply and taxes over the life-cycle. Blundell et al. (2016a) find that insurance via progressive taxation makes up a sizable contribution (\(~11\%\)) to insurance of permanent wage shocks, but other forces, most prominently (family) labor supply, dominate. For Norwegian data analyzed in Blundell et al. (2015) the riskiness of earnings is strongly attenuated by the tax and transfer system, especially for those with lower education, who experience roughly 20\% less impact of a permanent shock of one standard deviation on annual disposable compared to annual market income. Heathcote et al. (2014) simply aim to calculate the overall insurance provided by the aforementioned mechanisms. They do, however, find that own labor supply dampens the effect persistent shocks have on consumption by roughly 15\%. Heathcote et al. (2017a) go a step beyond the description of insurance mechanisms and provide a closed-form expression for social welfare, which crucially depends on the riskiness of earnings, and ultimately characterize the progressivity of the optimal income tax.

In the current paper I take a closer look at the descriptive side of insuring income fluctuations through progressive taxation. I evaluate how this insurance channel has been shaped by the policy maker. I do not provide a closed-form expression for social welfare, as Heathcote et al. (2017a) do, and am therefore silent on optimal policy. It is certainly the case that the policy maker should adopt some consistent stance on how to treat the trade-off between providing insurance and diminishing incentives for work. Therefore, I document the insurance effect of the chosen tax policy and its effect on net earnings risk.

3 The Model

The life-cycle behavior of individuals is described in the following. I give a detailed treatment of the way the tax and transfer system is modeled in sections 4-6.2.

The Life-Cycle Problem Life-cycle optimization by the individual proceeds by maximization of the sum of discounted in-period utilities from \(t_0\) to \(T\). I omit an individual specific subscript.
\[
\max_{c_t, h_t} E_{t_0} \left[ \sum_{t=t_0}^{T} \rho^{t-t_0} v(c_t, h_t, b_t) \right],
\]

where \( v \) is the in-period utility function taking consumption \( c_t \), hours of work \( h_t \), and taste-shifters \( b_t \) as arguments. \( \rho \) is the discount factor. I specify the taste shifter \( b_t = \exp(z\Xi_t - \upsilon_t) \). \( z\Xi_t \) is a linear combination of a set of personal characteristics. \( \upsilon_t \) accounts for the non-systematic variation of the taste shifter, which is assumed to be normally distributed and uncorrelated over time. The functional form of the in-period utility function is given by

\[
v(c_t, h_t, b_t) = \frac{c_t^{1-\theta}}{1-\theta} - b_t \frac{h_t^{1+\gamma}}{1+\gamma}, \quad \theta \geq 0, \gamma \geq 0,
\]

where \( \frac{1}{\theta} \) pins down the intertemporal elasticity of substitution with respect to consumption, while \( \frac{1}{\gamma} \) gives the Frisch-elasticity of labor supply. Thus, in-period utility is additively-separable and conforms to constant relative risk aversion (CRRA).

The intertemporal budget constraint is

\[
a_{t+1} = a_t + T_t(w_t h_t + N_t) - c_t,
\]

where \( a_t \) represents assets, \( r_t \) the real interest rate, \( N_t \) non-labor income, and \( T_t(\cdot) \) is the retention function that returns post-government/net income.

In the following I will examine interior solutions to this problem by estimating the associated labor supply equation derived from the first-order conditions.

**Uncertainty**  Wages evolve according to the equation

\[
\Delta \ln w_t = \alpha X_t + \Delta \omega_t,
\]

where \( X_t \) contains variables influencing human capital and \( \Delta \omega_t \) contains the innovation in the idiosyncratic shock processes.

It is assumed that the unobservable components determining wage growth can be decomposed into a permanent and a transitory process, which are chosen to be a random walk and a small-order moving average process (MA(1)). Thus, the dynamics of the idiosyncratic component of wages are described by the following set of equations:

---

3 Accordingly, the model features an incomplete capital market.

4 Choosing this process is standard in the extant literature and has favorable properties as far as identification is concerned.
\( \omega_{it} = p_{it} + \tau_{it} \) \hfill (5)

\( p_{it} = p_{it-1} + \zeta_{it} \)

\( \tau_{it} = \theta \epsilon_{it-1} + \epsilon_{it} \)

\( \zeta_{it} \sim N(0, \sigma^2_{\zeta,it}), \quad \epsilon_{it} \sim N(0, \sigma^2_{\epsilon,it}) \)

\( \mathbb{E}[\zeta_{it}\zeta_{-i} - l] = 0, \quad \mathbb{E}[\epsilon_{it}\epsilon_{-i} - l] = 0 \quad \forall l \in \mathbb{Z}_{\neq 0} \)

\( p_{it} \) is the permanent component and \( \tau_{it} \) the transitory. While \( p_{it} \) and \( \tau_{it} \) exhibit serial correlation, their innovations \( \zeta_{it} \) and \( \epsilon_{it} \) do not. With this error structure I can proceed to estimate the shock variances using the autocovariance moments of wages. The identification of the individual shock variances from year to year is discussed in 7.2 and in appendix 10.6. Suffice it to say that I need to impose an assumption regarding the initial conditions of the transitory shocks variances. In the empirical implementation I choose the zeroth and first transitory variance to have the same value. I determine the propagation of the shocks to hours and ultimately earnings in section 6.2.

**Empirical Implementation** Using the first order condition for labor supply and the wage equation in log changes, I can estimate residuals for both quantities. I assume that the levels of wages, hours and earnings are measured with error. I give my treatment of measurement error of in appendix 10.1. Using the second moments of the residuals from the labor supply and wage growth estimation I can recover the wage shock variances as well as the relevant parameters to determine shock pass-through to earnings. The roadblock in the way of the first step is how to model the function \( T(\cdot) \) in a way that makes the first step tractable. I will lay out the theoretical issues regarding this issue in section 4 and let the implementation follow in sections 5 and 6.

### 4 Approximation of the Tax and Transfer System

In brief, to proceed with the labor supply estimation, I need to approximate the retention function. The retention function \( T(\cdot) \) takes as inputs gross income as well as characteristics of the tax unit and returns net income. The function \( T(\cdot) \) is nonlinear, not continuous, and therefore non-differentiable. While it is not an issue to model the dependence on characteristics of the tax unit, it is not tractable to choose a non-continuous, non-differentiable retention function. When one proceeds with the first-order approach to labor supply estimation, while also approximating the life-time budget constraint, it turns out to be very advantageous to choose a power-function approximation of the tax

---

5See Blundell and MaCurdy (1999) and Keane (2011) for surveys on labor supply estimation and which approaches exist to implement taxation. Whether using an approximation is sensible, solely rests on how well the approximated post-government income corresponds to its observed counterpart.

6In this paper I will neither make an explicit distinction between pre-government and gross income nor between post-government and net income; I will define the terms in section 5.
function, as the resulting structural equations will be linear in log-space and the model becomes tractable.\footnote{See Blundell et al. (2016a) and Blundell et al. (2016b).}

There are three objectives for this theoretical section and the empirical section 6: First, in this section I want to introduce the commonly chosen power function approximation of the retention function and its connection to a common measure of progressivity of the tax system. Further, I will show that, when one introduces an explicit distinction between gross and taxable income in the retention function, the deductions determining this relationship may or may not have an impact on labor supply decisions. This depends on whether the function giving taxable income is nonlinear.

Second, in section 6 I want to provide accuracy measures of this approximation along the distribution of gross income. It is common to refer to goodness of fit measures, like the $R^2$ from the regression of log net on log gross or log taxable income while imposing the power-function. Implicitly, the reported $R^2$ uses a linear model for the tax function to calculate the total sum of squares. In some situations this comparison is certainly enlightening, however, one can better assess the immediate performance by directly inspecting the deviations between fitted and observed values. Therefore, I will generally rely on the root mean square error (RMSE) to evaluate fit. I will also calculate the implied fiscal gap arising from the approximation approach, so that I may gauge the effect on the government budget that the approximation will entail.

Third, I will explore the relationship between crucial parameters of the tax system, like the top tax rate and the size of the standard deduction, and the shape of the approximation function, in particular the parameter governing progressivity. This is crucial to detect the impact of a policy change on quantities like labor supply and the insurance provided through taxation.

Section 6 will also deal with the last steps toward tractability by extending the approximation for nonlabor income and allowing time-differencing.

### 4.1 The Power Function Approximation

Choosing a power function to represent the retention function was popularized by Feldstein (1969) and recent applications in the related literature include Kaplan (2012), Blundell et al. (2016a), Heathcote et al. (2014), and Heathcote et al. (2017a). The relationship between gross and net income is given by

$$T(y_{i,t}) \approx \chi y_{i,t}^{1-\tau}. \quad (6)$$

In this highly simplified version the parameters $\chi$ and $\tau$ are neither individual- nor time-specific. Accordingly such an approximation will miss out on much of the variation that is driven by i) the differences in the assessment criteria that apply to the particular tax unit, e.g. whether the tax unit consists of a couple filing jointly or a single, and ii) the differences in the relevant parameters of the tax code over time, e.g. the top tax rate.
For a single cross-section $t$ it is possible to introduce tax-unit specific variation in the parameters, by letting

$$T(y_{i,t}) \approx \chi_i (y_i)^{1-\tau_i},$$

and permitting individual variation through a parametric form such that

$$\chi_i = f\chi(Z_i),$$
$$\tau_i = f\tau(Z_i),$$

where the functions $f\chi$ and $f\tau$ may, for example, be linear in the tax unit’s characteristics $Z_i$. However, this type of modeling is incompatible with the structural labor supply estimation pursued later. Further, the focus of the current paper is on how the policy maker shapes individuals’ exposure to earnings risk. Hence, I will fix the parameters in the cross-section, but allow variation over time, so that

$$T_t(y_{i,t}) \approx \chi_t y_{i,t}^{1-\tau_t}. \quad (7)$$

To investigate the relationship between changes in the tax system and changes in the parameters of the approximation, I will need to isolate the mechanical effect that changes in the tax system have on the parameters of the approximation. I pursue the following three-step process to derive the mechanical effect of a change in the tax code on $\tau$: 1) For a given cross-section $t$ I derive the baseline estimate $\tau_t$. 2) I change the tax system in one relevant variable in a microsimulation program, e.g. increase the top marginal tax rate by one percent for the year $t$, run the counterfactual tax simulation and find the new estimate $\tau_t^c$. 3) Finally, I determine the percentage change in the parameter $\frac{\tau_t^c - \tau_t}{\tau_t}$ to determine the elasticity of the approximation parameters with respect to the tax parameters, i.e. the mechanical effect. I use NBER’s taxsim tax-calculator, which allows for only a couple of the tax code variations. I have chosen to restrict my attention to two parameters: the standard deduction and the top tax rate. The choice is motivated by the impact these two parameters have on the estimated tax function and their contested status in public policy debates. As described above, I will vary these parameters by one percentage point and calculate the impact on the parameters of the approximation. This exercise is detailed in section 6.1.3.

The above strategy does not come without drawbacks. Since I calculate the approximation using only the sample observed in this particular cross-section, there are bound to be characteristics of the sample that drive the approximation results.

### 4.2 The Measure of Progressivity

The choice of the power function to approximate the retention function entails a convenient link with a crucial metric discussed in the economics of taxation: the residual income progressivity or – in more intuitive terms – the elasticity of after-tax in-

---

8Estimation of the Frisch elasticity would have to be performed per group leading to potential power issues in estimation.
come with respect to gross income. Residual income progressivity can be expressed as
\[
\frac{\partial(y - T(y))}{\partial y} \frac{y}{y - T(y)} = \frac{1 - \partial T(y)/\partial y}{1 - T(y)/y},
\]
where \(T(\cdot)\) gives the tax liability. According to Jakobsson (1976), when comparing two tax schedules \(T_1\) and \(T_2\), one is more progressive than the other, when
\[
1 - \frac{\partial T_1(y)/\partial y}{1 - T_1(y)/y} < 1 - \frac{\partial T_2(y)/\partial y}{1 - T_2(y)/y} \quad \forall y.
\]
(8)

One implication of Jakobsson’s theorem is that for a progressive tax schedule
\[
\frac{\partial T(y)}{\partial y} > \frac{T(y)}{y} \quad \forall y.
\]
9
In the case of the power function approximation, \(\frac{\partial(y - T(y))}{\partial y} - \frac{y}{y - T(y)} = 1 - \tau\).

Further, the progressivity parameter \(1 - \tau\) directly impacts labor supply decisions, as shown in section 4.3, and the extent of insurance offered by taxation.

There is some ambiguity in the literature about how to measure the progressivity parameter. Specifically, it is possible to either measure progressivity by approximating net income based on gross income or net income minus deductions on taxable income; the latter being called statutory progressivity.\(^{10}\) I am interested in the parameters of the tax-function that affect behavior, which, in my setting, are the parameters affecting the choice of hours of work. First, I will alter the notation slightly to examine the relationship between effective and statutory marginal tax rates. I denote the statutory tax function by \(T_s(\cdot)\) and by \(\tilde{y}(y)\) the function mapping from gross to taxable income. The tax liability is given by the composition of both functions \(T_s(\tilde{y}(y))\). The statutory marginal tax rate is given by \(T'_s = \frac{\partial T_s}{\partial \tilde{y}}\), while the effective marginal tax rate is \(T'_e = \frac{\partial T_s}{\partial \tilde{y}} \frac{\partial \tilde{y}}{\partial y}\). This implies that statutory marginal rates bound effective marginal rates (weakly) from above, i.e.\(^{11}\)
\[
\frac{\partial T_s}{\partial \tilde{y}} \frac{\partial \tilde{y}}{\partial y} \leq \frac{\partial T_s}{\partial \tilde{y}}.
\]
(9)

This is evident, because the above expression holds with equality only when \(\frac{\partial \tilde{y}}{\partial y} = 1\). Whether this is the case is an empirical issue discussed in section 6.1.1. However, it is reasonable to expect that the effective tax function will be less progressive than the statutory one.

In the next section I use a static variant of the model in section 3 to fix ideas about how the retention function and its parameters affect labor supply decisions.

4.3 Modeling the Retention Function

Taking the model of section 3 while suppressing the time indices and setting non-labor income to zero, the first order condition for hours of work will take the form,

\[
h^\gamma = \frac{1}{b} \lambda T'(wh)w,
\]
(10)

\(T(\cdot)\) is the function giving net income, so I can replace it with \(y - T_s(\tilde{y}(y))\). Notice that, depending on whether the function \(T_s(\cdot)\) is nonlinear deductions will enter (10) and

\(^{10}\)Accordingly, when we approximate progressive tax schedules with continuous and differentiable functions, this implies that the approximation function has to be strictly convex. For intuition why this must be the case, construct the limiting case where \(T(\cdot)\) is linear and therefore \(\partial T(y)/\partial y = T(y)/y \forall y\).

\(^{11}\)See Blundell et al. (2016a) and Heathcote et al. (2017a) for illustrations of the two competing methods.

\(^{9}\)This statement only holds when \(\frac{\partial \tilde{y}}{\partial y} \leq 1 \forall y\), as one would reasonably anticipate.
therefore influence the individuals’ decision on labor supply. This is clearly the case when we consider a progressive tax system. Accordingly, in this setting I find it to be sensible to let the function giving taxable income also directly depend on gross income to retain tractability. A straightforward way to model both the tax and taxable income based on gross income is,

\[ T_s(\tilde{y}) = y - \tilde{\chi}y^{1-\tilde{\tau}}, \]
\[ \tilde{y}(y) = \kappa y^{1-\iota}. \]  (11)

This choice also conforms with the structure of the model of section 3 regarding consumption, since there is only one composite consumption good and no other good relevant for deductions. I illustrate more generally how deductions affect labor supply decisions in appendix 10.4, where I introduce a separate composite consumption good which can be deducted from gross income.\(^{12}\) The indication of that model, however, is the same as the one of equation (10): as long as the retention function possesses some nonlinearity, deductions do influence the first order condition for labor supply. With the approximation chosen, I resolve \(T'\) in (10) to obtain

\[ h^\gamma = \frac{1}{\kappa(1-\tilde{\tau})}\tilde{\chi}\left(\kappa(wh)^{1-\iota}\right)^{-\tilde{\tau}}\kappa(1-\iota)(wh)^{-\iota} w. \]  (12)

Applying logs and rearranging terms,

\[
\ln h = \frac{(1-\tilde{\tau})(1-\iota)\ln w - \ln b + \ln \lambda + \ln((1-\tilde{\tau})\tilde{\chi}) + \ln(\kappa(1-\iota) - \tilde{\tau})\ln \kappa}{\gamma + \tilde{\tau}(1-\iota) + \iota}.
\]  (13)

Which implies that the relevant tax-modified \(\lambda\)-constant elasticity, the analogue to the Frisch, is given by \(\frac{(1-\tilde{\tau})(1-\iota)}{\gamma + \tilde{\tau}(1-\iota)}\). Accordingly, with this specification for the retention function, not only the parameter of the statutory system \(\tilde{\tau}\), but also the parameter of the taxable income function \(\iota\) influence the individual’s hours-response to changes in the wage.

With these results, I can make some further observations. First, it follows from (11) that there exists a direct mapping from gross to net income, which is also a power function. Accordingly, whether I estimate both equations or just one power function approximation from gross to net income results in the same predictions for net income. I define \((1-\iota)(1-\tilde{\tau}) := 1-\tau\) and \(\tilde{\chi}\kappa^{1-\tilde{\tau}} := \chi\), so that the tax is given by

\[ T(y) = y - \chi y^{1-\tau}. \]  (14)

\(^{12}\)This approximation of net income and the tax system is different from the model of net income in Heathcote et al. (2017b), where deductions would appear as an extra additive term in the labor supply equation. This also implies that my statutory tax function is not comparable with their statutory tax function.
This is very convenient if only observations on net and gross income are available in the data or microsimulation of the tax and benefit system is not feasible.

Second, this also has an implication for how progressivity is measured. It becomes clear that progressivity of the entire system can be summarized as,

\[
\frac{\partial (y - T_s(\tilde{y}(y)))}{\partial y} \cdot \frac{y}{y - T_s(\tilde{y}(y))} = \tilde{\chi}(1 - \tilde{\tau}) (\kappa y^{1 - \iota})^{-\tilde{\tau}} \kappa (1 - \iota) y^{-\iota} \\
= (1 - \tilde{\tau})(1 - \iota) \tilde{\chi} \kappa^{1 - \tau} y^{-\iota - \tau + \tilde{\tau}} \\
= (1 - \tilde{\tau})(1 - \iota).
\]

Accordingly, when \( \iota \) increases, progressivity of the tax system increases.

Third, it is important to note that the chosen model implies no impact on progressivity when \( \iota = 0 \). In this case the overall progressivity \( \tau \) could be recovered both from the mapping between gross and net income, but also taxable and net income.

This final issue of whether a power or an affine function best represents the mapping from gross to taxable income is an empirical issue, which I will tackle in section 6 after introducing the data source in the next section.

5 Data

For all empirical exercises I use 9 waves of the Panel Study of Income Dynamics (PSID) coming from the survey years 1999 to 2015. The data are collected biennially and with reference to the previous year, so that the respective observed years are 1998 to 2014. I restrict the sample to households that either have a married or single head of household. Further, I follow Heathcote et al. (2017a) in restricting the sample to households with heads in prime working age, namely 25 to 60, and only keeping households with the head earning at least the equivalent of part-time work at the minimum wage. Finally, I drop observations with less than 260 and more than 4000 yearly hours. These restrictions allow me to focus on the population unambiguously participating in the labor market. Monetary variables are deflated to the base year 1998 using the CPI-U. Sample statistics of relevant variables for the final panel are presented in the reference year 2000 are presented in table 1.

Tax Variables I calculate all taxation variables using the taxsim tax calculator provided by the National Bureau of Economic Research (see Feenberg and Coutts (1993)). The preparation of the data before applying the tax calculator is done by adapting files provided by Kimberlin et al. (2016) accessible on the NBER taxsim webpage. Gross income includes all labor and non-labor income of the household, like interest, dividend, and rent income, which is marked as \( \text{PSID} (2015) \) and see PSID (2017) for an introduction to the PSID. 13

13 This amounts to an hourly wage of $5.15 deflated to 1998 times 1000 yearly hours. 14

14 Complete sample statistics for every year are presented in appendix 10.2. The year 2000 is fairly representative, although the sample size increases in 2002 and 2008 to a level slightly above 5100 and 5400 observations respectively. 15
### Table 1. Sample Statistics in Year 2000

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>hours</th>
<th>wage</th>
<th>years of education</th>
<th>num. of children</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>40.80</td>
<td>2164.25</td>
<td>19.88</td>
<td>13.22</td>
<td>1.08</td>
</tr>
<tr>
<td>sd</td>
<td>9.16</td>
<td>554.98</td>
<td>21.39</td>
<td>2.47</td>
<td>1.19</td>
</tr>
<tr>
<td>min</td>
<td>25</td>
<td>260</td>
<td>2.15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>max</td>
<td>60</td>
<td>4000</td>
<td>447.29</td>
<td>17</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>taxation</th>
<th>gross</th>
<th>tax</th>
<th>net</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>64508</td>
<td>18921</td>
<td>50411</td>
</tr>
<tr>
<td>sd</td>
<td>71911</td>
<td>29591</td>
<td>46437</td>
</tr>
<tr>
<td>min</td>
<td>5648</td>
<td>-3481</td>
<td>5231</td>
</tr>
<tr>
<td>max</td>
<td>2026291</td>
<td>905915</td>
<td>1145453</td>
</tr>
</tbody>
</table>

**Note:** Own calculation based on PSID (2015). All statistics are unweighted.

plus half of FICA.\(^\text{16}\) Deductions are calculated using PSID data on mortgage interest, medical expense and charitable giving deductions, while deductions stemming from state taxes are calculated by taxsim. The tax liability contains federal, state and FICA tax. Postgovernment or net income is gross income minus the tax liability plus transfers and half of FICA (employer’s share). Transfers, except for the EITC, which is calculated by taxsim, are recorded in the PSID data, i.e. TANF, social security, unemployment benefits, workers’ compensation and veterans’ pensions.

### 6 Estimating the Tax Function Approximation

In this section I address two broad sets of issues: First, I cover the questions raised in the previous section and explore the fit of the approximation, how relevant deductions are in shaping progressivity, and the policy maker’s influence over measured progressivity. Results on these issues apply across the board when a power function is used to approximate the retention function and are not specific to my application. Second, I implement a further modification to the approximation to allow for nonlabor income and I examine the issue of time-differencing with the power function present in the labor supply equation. These are tractability issues that are specific to this paper only.

#### 6.1 General Results

##### 6.1.1 Estimating the Taxable Income Function

I estimate both a linear (\(\iota = 0\)) and a power function for the relationship between gross and taxable income. If the relationship is linear, then deductions do not affect progressivity. Table 2 shows the estimated values for the parameters \(\iota\) and \(\kappa\) over time in the unrestricted model and in the restricted model (\(\iota = 0\)). Table 2 further shows a likelihood ratio test of the restricted model and a comparison of the root mean square error (RMSE) of the

\(^{16}\text{FICA or FICA tax is a payroll tax to fund social security and medicare payments.}\)
two models. $\iota$ is slightly below zero and so the exponent in the model is greater than one, making the estimated function convex. This indicates that retention of taxable from gross income rises over the distribution of gross income. However, $\frac{\partial \hat{y}(y)}{\partial y} < 0$ over the relevant range, so that overall progressivity of the tax system is lower due to deductions, as hypothesized in section 4.2. The values of $\iota$ exhibit some variation over time, but group pretty tightly in the range of -0.05 to -0.09. In the restricted model, with $\iota$ set to zero, the values for $\kappa$ lie in the range of 0.8 to 0.9 and are much larger than in the unrestricted model. The simple linear model implies that roughly 85 percent of a tax unit’s income is taxable. The likelihood ratio tests, however, show that the null hypothesis of model equivalence is easily rejected in all years, indicating that the taxable income function is nonlinear. This is also evidenced by the consistently lower RMSE for the unrestricted model. In relative differences the unrestricted model delivers a roughly 35% lower RMSE over the range of years when the restricted model is the base. Therefore, I reject the linear model of taxable income.

To visualize the fit I plot predicted versus actual values of both models for the year 2000 in figure 1. It is clear that the unrestricted model fits the mass at the lower end of the gross income distribution better than the restricted model, while also being able to curve up and fit values at the tail of the distribution. However, both functions have to start in the origin and fail to fit values with positive gross and zero taxable income well.

Table 2. Taxable Income Function

<table>
<thead>
<tr>
<th>Year</th>
<th>$\kappa$</th>
<th>$1 - \iota$</th>
<th>$\kappa$</th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>LM-Test</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>0.263</td>
<td>1.091</td>
<td>0.813</td>
<td>76401</td>
<td>0</td>
<td>7773</td>
<td>11652</td>
</tr>
<tr>
<td>2000</td>
<td>0.245</td>
<td>1.096</td>
<td>0.827</td>
<td>82065</td>
<td>0</td>
<td>8342</td>
<td>12794</td>
</tr>
<tr>
<td>2002</td>
<td>0.373</td>
<td>1.063</td>
<td>0.854</td>
<td>83267</td>
<td>0</td>
<td>8283</td>
<td>12642</td>
</tr>
<tr>
<td>2004</td>
<td>0.339</td>
<td>1.069</td>
<td>0.903</td>
<td>109172</td>
<td>0</td>
<td>10427</td>
<td>18149</td>
</tr>
<tr>
<td>2006</td>
<td>0.310</td>
<td>1.076</td>
<td>0.829</td>
<td>80580</td>
<td>0</td>
<td>8605</td>
<td>12836</td>
</tr>
<tr>
<td>2008</td>
<td>0.388</td>
<td>1.058</td>
<td>0.846</td>
<td>86231</td>
<td>0</td>
<td>8819</td>
<td>13473</td>
</tr>
<tr>
<td>2010</td>
<td>0.317</td>
<td>1.073</td>
<td>0.791</td>
<td>50267</td>
<td>0</td>
<td>8313</td>
<td>10673</td>
</tr>
<tr>
<td>2012</td>
<td>0.388</td>
<td>1.058</td>
<td>0.854</td>
<td>91930</td>
<td>0</td>
<td>7595</td>
<td>12227</td>
</tr>
<tr>
<td>2014</td>
<td>0.325</td>
<td>1.073</td>
<td>0.836</td>
<td>99186</td>
<td>0</td>
<td>7613</td>
<td>12280</td>
</tr>
</tbody>
</table>

Note: Own calculation based on PSID (2015). Restricted ($\iota = 0$) and unrestricted model estimated using nonlinear least squares. Estimation was performed using cross-sectional frequency weights. The relative difference of RMSEs is calculated in the following way: $\frac{\text{RMSE}_{\text{unres.}} - \text{RMSE}_{\text{res.}}}{\text{RMSE}_{\text{res.}}}$. 

6.1.2 Estimating the Retention Function

I now estimate the power function mapping for the retention function in two different ways: First, I estimate the mapping from taxable to net income (partial retention function) and then the combined relationship, i.e. the mapping from gross to net income (complete retention function). A visual illustration of the two approaches is given by figure 2. Qualitatively, both models fit a globally concave function, but the implied estimates for the progressivity parameter are clearly different. This can further be verified from table
Figure 1. Taxable Income Model Fit in 2000

Note: Own calculation based on PSID (2015). The graph plots gross income against taxable income. Light gray dots indicate observed values, black dots are predictions based on the unrestricted model and gray dots are predictions from the restricted model. Gross income above $300000 not shown.

I find that the progressivity parameter of the partial retention function $1 - \tilde{\tau}$ is roughly of the size 0.8 to 0.88. However, here I need to stress again that these estimates are not directly comparable with those in Heathcote et al. (2017a), since the authors model the relationship between net income minus deductions and taxable income.

The central result of this section is that the progressivity parameter of the complete retention function is in the range of 0.89 to 0.95. These estimates are reassuringly close to the estimates found in Blundell et al. (2016a) and to estimates in Gouveia and Strauss (1994). Blundell et al. (2016a) estimate a progressivity parameter of roughly 0.9 and Gouveia and Strauss (1994) arrive at values between 0.92 and 0.95 for the period from 1979 to 1989. These estimates diverge from the one provided by Heathcote et al. (2017a). Their estimate suggests more progressivity, with $1 - \tau = 0.82$. Since this is ultimately the parameter that will adjust the Frisch-elasticity of labor supply the comparison is appropriate. I comment on the discrepancy below.

I give a more intuitive description of the fit for the complete model by calculating the fiscal gap implied by the gross-to-net approximation. This fiscal gap is the difference between the predicted and the observed tax liability based on taxsim ($T - \hat{T}$). I plot the fiscal gap over the distribution of gross income in figure 3 for the year 2000. As one would expect, there are a lot of observations in the left tail of the income distribution up to

---

17Blundell et al. (2016a) do not use the taxsim tax-calculator, but have rather programmed their own.
Figure 2. Partial and Complete Retention Function Fit in 2000

Note: Own calculation based on PSID (2015). Complete model shown in the upper and partial in the lower panel. The graphs plot gross/taxable income against net income. Light gray dots indicate observed values, black dots are model predictions. Gross income above $300000 not shown.

roughly $100000 of gross income that achieve an acceptable fit. This is also evidenced by the fit of the LOWESS-line\textsuperscript{18} centering around zero. However, many observations do have overpredicted tax liabilities, which is clear from a somewhat strong spread into negative values at the very low end of the income distribution. Up to roughly $200000 residuals are centered on zero. Beyond this threshold the LOWESS-line bends up, which indicates that tax liabilities are underpredicted.

Further, to illustrate the impact on the budget of the state, I also calculate the cumulative and the average fiscal gap in every year. They are shown in the third- and second-to-last columns of table 3. Overall, the approximation imposes higher tax liabilities in every

\textsuperscript{18}The acronym LOWESS stands for locally weighted scatterplot smoothing, a local linear regression method.

year compared to the taxsim values. The pattern over the years is not uniform. There are years with larger fiscal gaps, like 2004 and 2008, but the average overestimation of the fiscal gap is roughly $64 million. This number can be more easily interpreted when one considers the adjacent column, which contains the average fiscal burden. On average each tax unit must pay $650 more than under the actual tax schedule. The average tax liability over all years is roughly $17000, so that the approximation would impose an increase of 3.82% of the average tax liability.

### Table 3. Retention Function

<table>
<thead>
<tr>
<th>Year</th>
<th>Partial Model $\bar{\chi}$</th>
<th>Partial Model $1 - \bar{\tau}$</th>
<th>Complete Model $\chi$</th>
<th>Complete Model $1 - \tau$</th>
<th>Absolute Fiscal Gap cumul. avg.</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>8.905</td>
<td>0.806</td>
<td>2.456</td>
<td>0.897</td>
<td>-29588838 -314</td>
<td>4535</td>
</tr>
<tr>
<td>2000</td>
<td>9.384</td>
<td>0.802</td>
<td>2.610</td>
<td>0.893</td>
<td>-22463340 -234</td>
<td>4878</td>
</tr>
<tr>
<td>2002</td>
<td>5.854</td>
<td>0.842</td>
<td>2.298</td>
<td>0.903</td>
<td>-54895752 -558</td>
<td>5104</td>
</tr>
<tr>
<td>2004</td>
<td>3.519</td>
<td>0.888</td>
<td>1.373</td>
<td>0.949</td>
<td>-164579872 -1671</td>
<td>5133</td>
</tr>
<tr>
<td>2006</td>
<td>6.526</td>
<td>0.836</td>
<td>2.121</td>
<td>0.912</td>
<td>-49338128 -490</td>
<td>5191</td>
</tr>
<tr>
<td>2008</td>
<td>4.086</td>
<td>0.876</td>
<td>1.692</td>
<td>0.931</td>
<td>-971311440 -955</td>
<td>5415</td>
</tr>
<tr>
<td>2010</td>
<td>6.892</td>
<td>0.833</td>
<td>2.318</td>
<td>0.905</td>
<td>-36420912 -362</td>
<td>5141</td>
</tr>
<tr>
<td>2012</td>
<td>4.128</td>
<td>0.876</td>
<td>1.775</td>
<td>0.928</td>
<td>-80230176 -831</td>
<td>5225</td>
</tr>
<tr>
<td>2014</td>
<td>6.973</td>
<td>0.829</td>
<td>2.401</td>
<td>0.901</td>
<td>-45120448 -435</td>
<td>5346</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-64418767 -650</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Own calculation based on PSID (2015). Both models estimated using nonlinear least squares. The fiscal gap is calculated using predicted tax liabilities from taxsim and the complete model. Estimation and ancillary calculations were performed using cross-sectional frequency weights.*

### Relationship to the Estimates of Heathcote et al. (2017a)

Finally, I want to explain why Heathcote et al. (2017a) find a different estimate for the progressivity parameter $1 - \tau$. The clearest conceptual distinction to point out is that Heathcote et al. (2017a) measure statutory progressivity by regressing the log of net income minus deductions on the log of taxable income.\(^{19}\) Even though the conceptual difference is large, empirically it appears to matter little. Rather, it is of crucial importance whether one estimates the model using nonlinear least-squares or with OLS after a log-transformation.

In appendix 10.5 I have run the log specification and the equivalent nonlinear least-squares specification using the data that Heathcote et al. (2017a) provide for replication. I estimate both models pooled over the period 2000-2006. When using nonlinear least-squares I arrive at a much larger value of about 0.93 instead of 0.82 for the progressivity parameter. When comparing the RMSE of the residuals in levels, the log specification performs worse than the nonlinear one; the RMSE is slightly more than twice as large. This is because the log specification implicitly puts less weight on observations at the top end of the taxable income distribution and overpredicts their tax liabilities.

The reason why the estimates differ lies in the specification and not in the conceptual difference between effective or statutory progressivity, as my “effective” estimate of $\tau$ is

---

\(^{19}\)In their model, deductions, at least at the margin, are fixed and do not react to income changes.
very close to the “statutory” $\tau$ found using nonlinear least-squares using their data.\footnote{It can further be argued that their approximation function only partly captures statutory progressivity. Heathcote et al. (2017a) do include transfers in their measure of net income and therefore count transfers as part of the statutory system. However, by regressing the log of net income minus deductions on the log of taxable income, individuals with zero taxable income with positive net income achieved with transfers cannot affect the progressivity estimate as they are excluded from the regression. The same would hold if they estimated the relationship using nonlinear least-squares, since the function has to start in the origin. This problem does not arise when one directly maps from gross to net income, provided the sample is restricted to the active population.} So, although I am convinced that it is appropriate to model the effective progressivity following the arguments in section 4.3 and appendix 10.4, the importance of whether the effective or the statutory progressivity is modeled is ultimately of minor importance for the empirics.

6.1.3 The Policy Maker’s Influence over Progressivity

As discussed above, it is vital for policy making to have a measure of the impact a change in tax policy has on the progressivity parameters of the approximation. Therefore, I calculate the mechanical effect of policy changes on the parameters of the approximation determining progressivity.

I conduct two experiments using the non-standard options of the taxsim tax-calculator: 1) I increase the standard deduction available to the tax units by one percent. 2) I increase the top tax rate by one percent.

Following that, I calculate taxable and net income, and run the approximation procedures. I recover the new, counterfactual approximation parameters $\iota^c$ and $\tau^c$, and compute relative differences from the baseline in each year. These relative differences give the elas-
ticity of the approximation parameters with respect to a change in the tax function, i.e. the desired mechanical effect of a change in the tax function. Table 4 lists the effects on $\iota$ and on $\tau$ of the two scenarios.

Table 4. Elasticities of the Approximation Parameters

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>deduction</td>
<td>$\iota$</td>
<td>0.485</td>
<td>0.419</td>
<td>0.465</td>
<td>0.386</td>
</tr>
<tr>
<td></td>
<td>$\tau$</td>
<td>0.117</td>
<td>0.101</td>
<td>0.084</td>
<td>0.122</td>
</tr>
<tr>
<td>top tax rate</td>
<td>$\iota$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\tau$</td>
<td>1.380</td>
<td>1.424</td>
<td>1.725</td>
<td>2.730</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>scenario</th>
<th>2008</th>
<th>2010</th>
<th>2012</th>
<th>2014</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>deduction</td>
<td>$\iota$</td>
<td>0.499</td>
<td>0.815</td>
<td>0.579</td>
<td>0.564</td>
</tr>
<tr>
<td></td>
<td>$\tau$</td>
<td>0.103</td>
<td>0.135</td>
<td>0.113</td>
<td>0.107</td>
</tr>
<tr>
<td>top tax rate</td>
<td>$\iota$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\tau$</td>
<td>1.954</td>
<td>1.164</td>
<td>1.896</td>
<td>1.577</td>
</tr>
</tbody>
</table>

Note: Own calculation based on PSID (2015). Approximation models computed in the same way as in sections 6.1.1 and 6.1.2. Calculations performed using cross-sectional frequency weights. Scenario deduction corresponds to an increase of the standard deduction by one percent. Scenario top tax rate corresponds to an increase of the top tax rate by one percent.

The most noticeable feature of the results is that a change in the standard deduction affects both $\iota$ and $\tau$, but a change in the top tax rate does not affect $\iota$. But this is expected because a change in the top tax rate should not affect the mapping from gross to taxable.

In terms of the magnitude of the reactions, the picture is relatively uniform. A one percent increase in the standard deduction increases $\iota$ by half a percent. Since an increase in $\iota$ necessitates an increase in overall progressivity, $\tau$ rises by roughly a tenth of a percent. The increase of the top tax rate by one percent shows a much greater effect on overall progressivity. The change in $\tau$ in the mean is roughly 1.7 percent. However, the variability of the estimates is a lot higher in this scenario, which is likely driven by fluctuations in the size and the respective incomes of the group paying the top tax rate.

6.2 Tractability Results

6.2.1 Deriving the Labor Supply Equation

In the previous sections I have introduced the power function as a convenient approximation of the retention function. However, to make the approximation fully tractable with the first order approach of estimating labor supply, I need to make a further adjustment to allow for nonlabor income and time-differencing.

To allow for non-labor income $N_t$, I will adjust the approximation slightly and show that in terms of tractability of the first-order approach the adjustment has very desirable properties, while the cost in terms of fit is acceptable.

The life-cycle model is the same as described in section 3. I choose the following functional form for $T_t(\cdot)$:
\[ T_t(w_t h_t, N_t) \approx \tilde{\chi}_t (\kappa_t (w_t h_t)^{1-\tau_t})^{1-\tilde{\tau}_t} + \tilde{\chi}_t (\kappa_t (N_t)^{1-\tau_t})^{1-\tilde{\tau}_t} \]
\[ = \chi_t (w_t h_t)^{1-\tau_t} + \chi_t (N_t)^{1-\tau_t}, \quad (15) \]

so that the budget constraint takes the form,
\[
\frac{a_{t+1}}{(1+r_t)} = a_t + \chi_t (w_t h_t)^{1-\tau_t} + \chi_t (N_t)^{1-\tau_t} - c_t. \quad (16)
\]

The additional assumption imposed in contrast to the approximations in 4.1 is that one can separately apply the power function approximation to labor earnings and non-labor income and arrive at the same net income. I call this the additive approximation and will discuss how appropriate this alternate choice is in terms of fit in section 6.2.2. The first order condition for choosing labor supply in this setting is
\[
h_t^{\gamma} = b_t \lambda_t \chi_t (1-\tau_t) (w_t h_t)^{-\tau_t} w_t. \quad (17)
\]

To make the estimation in growth rates, which eliminates \( \lambda_t \), tractable, I posit the following lagged approximation:
\[
T_{t-1}(w_{t-1} h_{t-1}, N_{t-1}) \approx \tilde{\chi}_{t-1} (w_{t-1} h_{t-1})^{1-\tau_t} + \tilde{\chi}_{t-1} (N_{t-1})^{1-\tau_t}. \quad (18)
\]

Here the degree of progressivity is set to the value of the upcoming period and the parameter \( \tilde{\chi} \) can freely adjust to fit the period \( t-1 \) distribution of post-government income. Again, the question whether it is possible to make this substitution depends on the fit of the approximation, which I discuss in section 6.2.3.

Divide (17) by \( h_{t-1}^{\gamma} \), so that
\[
\left( \frac{h_t}{h_{t-1}} \right)^{\gamma} = \frac{b_{t-1}}{b_t} \frac{\lambda_t}{\tilde{\lambda}_{t-1}} \frac{\chi_t}{\tilde{\chi}_{t-1}} (1-\tau_t) \left( \frac{w_t}{w_{t-1}} \frac{h_t}{h_{t-1}} \right)^{-\tau_t} \frac{w_t}{w_{t-1}}. \quad (19)
\]

Taking logs I find that
\[
\gamma \Delta \ln h_t = \Delta \ln \lambda_t + (\ln \chi_t - \ln \tilde{\chi}_{t-1}) - \tau_t \Delta \ln h_t + (1-\tau_t) \Delta \ln w_t - \Delta \ln b_t
\]
\[
\Delta \ln h_t = \frac{1}{\gamma + \tau_t} [\Delta \ln \lambda_t + (\ln \chi_t - \ln \tilde{\chi}_{t-1}) + (1-\tau_t) \Delta \ln w_t - \Delta \ln b_t]
\]
\[
\Delta \ln h_t = \frac{1}{\gamma + \tau_t} [\text{cons} + \Delta \ln \lambda_t + (1-\tau_t) \Delta \ln w_t - \zeta \Delta \Xi_t + \Delta \nu_t]. \quad (20)
\]

The term \( \text{cons} \) contains all the terms not varying in the cross-section. Finally, I obtain the estimating equation for labor supply by resolving the expression for the intertemporal difference in the log marginal utility of wealth. Everything in that term, except for the innovations \( \eta_t \), are absorbed into the constant. I derive the approximation of the Euler equation that underlies this substitution in appendix 10.3. Then,
\[
\Delta \ln h_t \approx \frac{1}{\gamma + \tau_t} \left[ \text{const}_t + (1 - \tau_t) \Delta \ln w_t - \zeta \Delta \Xi_t + \Delta \nu_t + \eta_t \right].
\] (21)

This tax-modified Frisch elasticity can be estimated by IV techniques. Its estimation is the objective in section 7.1.

The innovations \( \eta_t \) are purely a function of the permanent wage shocks \( \zeta_{it} \). An approximation of the life-time budget constraint as described in Blundell et al. (2016a) reveals

\[
\eta_{it} = - \varphi_{it} (1 - \tau_t) \left( 1 + \frac{1 - \tau_t}{\gamma + \tau_t} \right) \zeta_{it}, \quad \varphi_{it} \sim LN (\mu_\varphi, \sigma_\varphi^2).
\] (22)

\( \varphi_{it} \) is a transmission parameter measuring how permanent wage shocks affect \( \eta_{it} \). It is individual specific because it is mainly driven by the ratio of wealth to total wealth including human wealth (Jessen and König, 2018). Further, the parameter is necessary to compute the Marshallian elasticity of labor supply. Since I do not use consumption or asset data, I choose a log-normal distribution with the underlying parameters \( \mu_\varphi \) and \( \sigma_\varphi \) to model its distribution, as \( \varphi_{it} \) can only take on non-negative values.

### 6.2.2 Additive Approximation

The main concern regarding the use of the additive formulation of the tax function approximation is whether it restricts the goodness of fit or whether it gravely alters the parameter estimates relevant for the calculation of progressivity. This is not the case. The additive formulation neither has a major impact on the parameter estimates for \( \iota \) and \( \tau \), nor does it gravely change the goodness of fit. Unfortunately it is not possible to conduct a direct, formal test because the models are not nested. However, it is obvious from the results in table 5 that the two approximations deliver very similar results in terms of coefficients and goodness of fit.

The table’s contents are reassuring: The parameters that determine progressivity, \( \iota \) and \( \tau \), are exceptionally close in almost every year. Despite the major alteration of the approximation, my results in terms of the effect of taxation on labor supply are bound to be comparable to the rest of the literature. The largest differences occur in 2014, where \( 1 - \iota \) is roughly one percent smaller and \( 1 - \tau \) roughly one percent larger in the additive model. The multiplicative parameters, \( \kappa \) and \( \chi \), show larger differences and more variability, but this will not impact the estimation of the Frisch elasticity.

Although goodness of fit in terms of the RMSE does not have a uniform pattern, it is very close in almost all years. In some years, like 1998 and 2008, the additive deduction model even dominates the original, while in 2004 and 2008 the additive, complete model dominates the original. The overall impression is that the original model has better fit in most years. However, the difference is slight: the biggest relative difference in the deduction model is recorded in 2002, where the RMSE is 2.6 percent larger in the additive model. The largest relative difference in the complete model also occurs in 2002; a difference of 6.7 percent. Generally, one can observe that relative differences in the RMSE in the complete model are larger, clustering in the range of 0.3 to 6.2 percent, excluding 2004 and 2008. In
sum, I can conclude that the additive model does almost as good of a job of fitting taxable and net income as the original model.

Table 5. Additive Tax Function

<table>
<thead>
<tr>
<th>Year</th>
<th>1 - τ</th>
<th>κ</th>
<th>1 - τ</th>
<th>κ</th>
<th>add.</th>
<th>orig.</th>
<th>rel. diff.</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>1.090</td>
<td>0.275</td>
<td>1.091</td>
<td>0.263</td>
<td>7692</td>
<td>7773</td>
<td>-0.010</td>
<td>4535</td>
</tr>
<tr>
<td>2000</td>
<td>1.094</td>
<td>0.263</td>
<td>1.096</td>
<td>0.245</td>
<td>8407</td>
<td>8342</td>
<td>0.008</td>
<td>4878</td>
</tr>
<tr>
<td>2002</td>
<td>1.056</td>
<td>0.414</td>
<td>1.062</td>
<td>0.373</td>
<td>8496</td>
<td>8283</td>
<td>0.026</td>
<td>5104</td>
</tr>
<tr>
<td>2004</td>
<td>1.071</td>
<td>0.344</td>
<td>1.069</td>
<td>0.339</td>
<td>10676</td>
<td>10427</td>
<td>0.024</td>
<td>5133</td>
</tr>
<tr>
<td>2006</td>
<td>1.069</td>
<td>0.351</td>
<td>1.076</td>
<td>0.310</td>
<td>8786</td>
<td>8605</td>
<td>0.021</td>
<td>5191</td>
</tr>
<tr>
<td>2008</td>
<td>1.057</td>
<td>0.400</td>
<td>1.057</td>
<td>0.388</td>
<td>8717</td>
<td>8819</td>
<td>-0.012</td>
<td>5145</td>
</tr>
<tr>
<td>2010</td>
<td>1.066</td>
<td>0.353</td>
<td>1.073</td>
<td>0.317</td>
<td>8329</td>
<td>8313</td>
<td>0.002</td>
<td>5141</td>
</tr>
<tr>
<td>2012</td>
<td>1.051</td>
<td>0.430</td>
<td>1.058</td>
<td>0.388</td>
<td>7652</td>
<td>7595</td>
<td>0.008</td>
<td>5225</td>
</tr>
<tr>
<td>2014</td>
<td>1.064</td>
<td>0.372</td>
<td>1.073</td>
<td>0.325</td>
<td>7783</td>
<td>7613</td>
<td>0.022</td>
<td>5346</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>1 - τ</th>
<th>χ</th>
<th>1 - τ</th>
<th>χ</th>
<th>add.</th>
<th>orig.</th>
<th>rel. diff.</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>0.899</td>
<td>2.277</td>
<td>0.897</td>
<td>2.456</td>
<td>6587</td>
<td>6582</td>
<td>0.001</td>
<td>4535</td>
</tr>
<tr>
<td>2000</td>
<td>0.895</td>
<td>2.402</td>
<td>0.893</td>
<td>2.610</td>
<td>13230</td>
<td>13222</td>
<td>0.001</td>
<td>4878</td>
</tr>
<tr>
<td>2002</td>
<td>0.911</td>
<td>2.007</td>
<td>0.903</td>
<td>2.298</td>
<td>4362</td>
<td>4089</td>
<td>0.067</td>
<td>5104</td>
</tr>
<tr>
<td>2004</td>
<td>0.946</td>
<td>1.379</td>
<td>0.949</td>
<td>1.373</td>
<td>6925</td>
<td>7177</td>
<td>-0.035</td>
<td>5133</td>
</tr>
<tr>
<td>2006</td>
<td>0.919</td>
<td>1.868</td>
<td>0.912</td>
<td>2.121</td>
<td>4754</td>
<td>4505</td>
<td>0.055</td>
<td>5191</td>
</tr>
<tr>
<td>2008</td>
<td>0.933</td>
<td>1.609</td>
<td>0.931</td>
<td>1.692</td>
<td>4589</td>
<td>4626</td>
<td>-0.008</td>
<td>5415</td>
</tr>
<tr>
<td>2010</td>
<td>0.913</td>
<td>2.031</td>
<td>0.905</td>
<td>2.318</td>
<td>4369</td>
<td>4245</td>
<td>0.029</td>
<td>5141</td>
</tr>
<tr>
<td>2012</td>
<td>0.936</td>
<td>1.566</td>
<td>0.928</td>
<td>1.775</td>
<td>4639</td>
<td>4485</td>
<td>0.034</td>
<td>5225</td>
</tr>
<tr>
<td>2014</td>
<td>0.911</td>
<td>2.048</td>
<td>0.901</td>
<td>2.401</td>
<td>4544</td>
<td>4391</td>
<td>0.035</td>
<td>5346</td>
</tr>
</tbody>
</table>

Note: Own calculation based on PSID (2015). Models estimated using nonlinear least squares. Estimation and ancillary calculations were performed using cross-sectional frequency weights. Relative difference between RMSEs calculated using this expression: \( \frac{RMSE^{add} - RMSE^{orig}}{RMSE^{orig}} \).

6.2.3 Approximation of Lagged Net Income

The final issue to resolve is the evaluation of the fit of the tax approximation when restricting the progressivity parameter \( 1 - τ \) in the approximation of the \( t - 1 \)-tax-system to the value estimated in \( t \). In this restricted model only \( χ \) can adjust. The corresponding unrestricted model is the additive, effective model shown in table 5. I present the RMSE of both the restricted and the unrestricted model in table 6 along with a linear one for comparison. The linear model is a natural benchmark, since it restricts the effect of taxation on labor supply to be nil.

At first glance table 6 shows that the restricted model generates noticeable differences in terms of fit in years like 2002 and 2012 compared to the unrestricted model. However, since changes in progressivity between years cannot be fully accounted for by letting only \( χ \)
adjust, this loss in fit is expected. When compared to the linear model, in which $\tau$ is set to zero, the restricted still compares favorably. The loss of fit is primarily an issue in the upper tail of the distribution. In figure 4 I plot both the restricted and the unrestricted model predictions against the observed values for the year 2002, which has the worst relative gap in RMSE. It is clear from the figure that the most severe deviations from the unrestricted model only occur after levels of roughly $200000$ gross income. Below this threshold both models make very similar predictions. While the approximation with the imposed leading progressivity parameter is certainly not preferable to the unrestricted model, it does not lead to a major reduction in terms of goodness of fit. Further, it still handily outperforms the linear model. Accordingly, imposing this relationship implicitly when I estimate the labor supply equation (21) is a justifiable sacrifice to attain tractability.

### Table 6. Fit of the Retention Function with Lead of Progressivity Parameter

<table>
<thead>
<tr>
<th>Year</th>
<th>restr.</th>
<th>unres.</th>
<th>linear</th>
<th>restr. vs. unres.</th>
<th>restr. vs. lin.</th>
<th>unres. vs. lin.</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>6587</td>
<td>6595</td>
<td>10159</td>
<td>0.001</td>
<td>-0.352</td>
<td>-0.351</td>
<td>4535</td>
</tr>
<tr>
<td>2000</td>
<td>13230</td>
<td>13293</td>
<td>15791</td>
<td>0.005</td>
<td>-0.162</td>
<td>-0.158</td>
<td>4878</td>
</tr>
<tr>
<td>2002</td>
<td>4362</td>
<td>6148</td>
<td>11837</td>
<td>0.409</td>
<td>-0.631</td>
<td>-0.481</td>
<td>5104</td>
</tr>
<tr>
<td>2004</td>
<td>6925</td>
<td>8300</td>
<td>11260</td>
<td>0.199</td>
<td>-0.385</td>
<td>-0.263</td>
<td>5133</td>
</tr>
<tr>
<td>2006</td>
<td>4754</td>
<td>4950</td>
<td>9845</td>
<td>0.041</td>
<td>-0.517</td>
<td>-0.497</td>
<td>5191</td>
</tr>
<tr>
<td>2008</td>
<td>4589</td>
<td>5359</td>
<td>10603</td>
<td>0.168</td>
<td>-0.567</td>
<td>-0.495</td>
<td>5415</td>
</tr>
<tr>
<td>2010</td>
<td>4369</td>
<td>4779</td>
<td>8699</td>
<td>0.094</td>
<td>-0.498</td>
<td>-0.451</td>
<td>5141</td>
</tr>
<tr>
<td>2012</td>
<td>4639</td>
<td>5995</td>
<td>10671</td>
<td>0.292</td>
<td>-0.565</td>
<td>-0.438</td>
<td>5225</td>
</tr>
</tbody>
</table>

*Note: Own calculation based on PSID (2015). Models estimated using nonlinear least squares. Estimation and ancillary calculations were performed using cross-sectional frequency weights. Relative difference between RMSEs calculated using these expression from left to right:*  
\[ \frac{\text{RMSE}_{\text{restr.}} - \text{RMSE}_{\text{unres.}}}{\text{RMSE}_{\text{unres.}}} \],  
\[ \frac{\text{RMSE}_{\text{restr.}} - \text{RMSE}_{\text{lin.}}}{\text{RMSE}_{\text{lin.}}} \],  
\[ \frac{\text{RMSE}_{\text{unres.}} - \text{RMSE}_{\text{lin.}}}{\text{RMSE}_{\text{lin.}}} \].

### 7 Results

I now turn to the empirical implementation: First, I estimate both the tax-modified and the unmodified Frisch elasticity of labor supply and the residuals of the wage and hours equations. Second, I estimate the permanent and transitory shock variances of the wage process and finally the Marshallian elasticity of labor supply.

#### 7.1 First Stage - Labor Supply

To obtain the residuals of the wage equation (4), I perform a first-differenced regression of the log of hourly wages on indicator variables for calendar years, states, industries, occupations, number of children in the household, and race. Further, I include a second order polynomial of years of education and its interaction with age. This last set of variables is also used as the set of excluded instruments for wages in the labor supply estimation.
Figure 4. Fit of Restricted and Unrestricted Model in 2002

Note: Own calculation based on PSID (2015). The graph plots gross income against net income. Light gray dots indicate observed values, black dots are predictions based on the unrestricted model and gray dots are predictions from the restricted model. Gross income above $300000 not shown.

Now I specify the observable portion of the taste-shifter $\Xi_t$ in the labor supply equation (21). Again, I include indicators for years, states, industries, occupations, number of children in the household, and race. I first estimate equation (21) pooled over all years to find an estimate of the average tax-modified Frisch elasticity. It is an average over the years because the tax-modified elasticity varies with $\tau$, which in turn changes year to year. I use the average progressivity parameter over the years to back out the unmodified Frisch elasticity $1/\gamma$, as it does not vary over the years. The average of $1 - \bar{\tau}$ over the years relevant for estimation (2000-2014) can be determined from the pooled estimation of the additive model of equation (15), which is $1 - \bar{\tau} = 0.925$.

I display the estimated tax-modified and the unmodified Frisch elasticity along with the average progressivity parameter in table 7.

Table 7. Regression Results

<table>
<thead>
<tr>
<th></th>
<th>$1 - \bar{\tau}$</th>
<th>$(1 - \bar{\tau})/(\gamma + \bar{\tau})$</th>
<th>$1/\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.925</td>
<td>0.469</td>
<td>0.528</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.071)</td>
<td></td>
</tr>
<tr>
<td>Kleibergen Paap F Stat.</td>
<td>64.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>30558</td>
<td>30558</td>
<td></td>
</tr>
</tbody>
</table>

Note: Own calculation based on PSID (2015). Robust standard errors in parentheses. $(1 - \bar{\tau})/\gamma + \bar{\tau}$ is the tax-modified Frisch elasticity and $1/\gamma$ is the regular Frisch-elasticity. Observations are person-years.

The estimated tax-modified Frisch elasticity of labor supply is about 0.47 and the unmodified elasticity is slightly larger with a point estimate of 0.528. Both are statistically significant at conventional levels and the F statistic according to Kleibergen and Paap
(2006) does not indicate a weak-instrument issue. That unmodified Frisch elasticities are larger than tax-modified ones is known in the literature and expected. The tax-modified elasticity gives the response to a transitory wage change before taxes, while the unmodified Frisch gives the response after-tax. The tax-modified elasticity is adjusted for the disincentive effect of the progressive tax system, which lowers individuals’ responsiveness to a change in the wage. My result for the unmodified Frisch elasticity is in line with the estimates presented in Heathcote et al. (2014) (0.462) and in Blundell et al. (2016a) (0.681). That the estimates agree fairly well is encouraging as both studies employed the same data source and similar portions of the data, but they did not use the same method to estimate the unmodified Frisch elasticity. Rather, both use the method of moments applied to residual variances to estimate it. Note, however, that in Blundell et al. (2016a) the second earner is explicitly considered and, therefore, the structural equations on which the estimation rests are quite distinct from mine.

The difference between the tax-modified and the unmodified elasticity is rather small and most likely not statistically significant in spite of the small standard errors. The unmodified elasticity is roughly four percent larger than the tax-modified counterpart. In comparison, it is roughly 17 percent larger for Blundell et al. (2016a). But even in that study the increase is most likely not statistically significant.

The residuals for the second stage of the estimation are obtained from estimating equation (21) year by year to account for the time-dependence of the tax-modified Frisch elasticity.

### 7.2 Second Stage - Wage Variance Process

I estimate the stochastic process for wages described in 5. I provide an example of the identification for the process with \( t=3 \) in appendix 10.6. As I discussed in section 3, it is necessary to set an initial condition for the transitory process in period zero, so I restrict the innovation variance for the transitory shock in period 0 to have the same value as in period 1, \( \sigma^2_{\epsilon,0} = \sigma^2_{\epsilon,1} \). Further, as shown in appendix 10.6, I can only identify all the parameters of interest up to period \( t-1 \) if \( t \) periods are available.

**Method of Moments** I estimate the wage process using the method of moments with a unit weighting matrix. Let the set of parameters of interest be denoted by \( \Sigma \), so that it contains all the 7 permanent shock variances \( \sigma^2_{\zeta} \), 7 transitory shock variances \( \sigma^2_{\epsilon} \), and the persistence parameter \( \theta \). Then the minimization program for the method of moments is given by

\[
\min_{\Sigma} \left[ m(\Sigma) - m^c \right]'I\left[ m(\Sigma) - m^c \right],
\]

where \( m(\Sigma) \) is the vector of theoretical autocovariance moments of \( \Delta w_{it} \) and \( m^c \) is the observed counterpart. By choosing the identity matrix \( I \) as the weighting matrix, I

---

\(^{21}\) Heathcote et al. (2017a) use the unmodified Frisch elasticity estimated in Heathcote et al. (2014) for the calibration of their model.
minimize the squared sum of deviations between the observed and theoretical moments. I

calculate standard errors for the parameters using the block bootstrap method. I draw

200 bootstrap replications of the data.

I present the estimates of the persistence parameter and the standard deviations of the
two shock types in table 8.

Table 8. Wage Process

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sigma_{\zeta,t})</td>
<td>0.1853</td>
<td>0.2011</td>
<td>0.1970</td>
<td>0.2194</td>
<td>0.2503</td>
<td>0.2503</td>
<td>0.2374</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0028)</td>
<td>(0.0018)</td>
<td>(0.0014)</td>
<td>(0.0014)</td>
<td>(0.0011)</td>
<td>(0.0011)</td>
<td>(0.0009)</td>
<td></td>
</tr>
<tr>
<td>(\sigma_{\epsilon,t})</td>
<td>0.1728</td>
<td>0.2688</td>
<td>0.1900</td>
<td>0.1955</td>
<td>0.1941</td>
<td>0.1372</td>
<td>0.0615</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0019)</td>
<td>(0.0016)</td>
<td>(0.0025)</td>
<td>(0.0023)</td>
<td>(0.0022)</td>
<td>(0.0018)</td>
<td>(0.0030)</td>
<td></td>
</tr>
<tr>
<td>(\theta)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.0329</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0076)</td>
</tr>
</tbody>
</table>

*Note: Own calculation based on PSID (2015). Bootstrap standard errors based on 200 replications in parentheses.*

The time trend of the standard deviations of the two shock types is striking. The size

of permanent shocks to the wage process increases steadily until the financial crisis of 2008
hits. In 2008 the size of permanent shocks is extraordinarily large and remains at this level
in 2010 with a slight recovery in 2012. To contrast, the increase in the standard deviation
of permanent shocks from 2000 to 2006 is roughly 14 percent, while the increase from 2006
to 2008 is, again, 14 percent.

Transitory shocks show a completely different intertemporal pattern. While the size

of the standard deviation is comparable to the permanent counterpart in the first couple of
years, transitory shock size almost halves in 2006. This downtrend is not heavily impacted
by the financial crisis of 2008. Even in 2012 transitory shock size appears to decrease
further rather than tend back to pre-crisis levels.

This pattern is interesting on its own as it suggests that permanent, partially uninsur-
able wage risk has increased for the active population and not just due to the impact of the
financial crisis. The pattern of rising partially uninsurable wage risk is also documented
in Table E1 of ?. Even though they do not estimate the process I have chosen, they do re-
port close analogues, namely the uninsurable, island-level shock variances. They estimate
their process until 2006 and report a standard deviation of 0.1236 in 2000 for these shocks
and 0.1378 in 2006 when frequency-adjusted. This implies an increase of the standard
deviation from 2000 to 2006 of about 11 percent, showing that the relative trend lines up
in both sets of results.

\(22\) Altonji and Segal (1996) show that the identity weighting matrix is generally preferable for the esti-
mation of autocovariance structures using panel data.

\(23\) This accounts for various issues that would otherwise affect the more conventional Delta-method stan-
dard errors, which include the use of estimates for the variance of the measurement error, heteroskedasticity
and serial correlation.

\(24\) Heathcote et al. (2014) assume a distinction between insurable and uninsurable shocks from the outset
and rationalize it using an island-structure for the shock process that agents face (Attanasio and Rıos-Rull,
2000). I add the two shock variances reported in ? for the two years that are pooled in my dataset to
account for the frequency difference.
The results in Blundell et al. (2016a) are harder to compare because there is explicit consideration of a secondary earner, shocks are allowed to be correlated across primary and secondary earner, variances are calculated with respect to the age group and not the calendar year, and their sample period runs from 1998 to 2008. Still, their average figure for the permanent shock standard deviation primary earners face is 0.1741 and, considering all these caveats, fairly close to my estimates for the pre-crisis period.

### 7.3 Marshallian Elasticity

Finally, I use the autocovariance moments of the hours residuals and the covariance moments with the wage residuals to estimate the parameters of $\varphi$, $\mu_\varphi$ and $\sigma^2_\varphi$.\(^{25}\) Consistent with the estimation procedure above, this is done by the method of moments choosing the identity matrix as the weighting matrix. The reader should note that the parameter $\varphi$ does depend on $\tau_t$ and so the estimation delivers values for the parameters $\mu_\varphi$ and $\sigma^2_\varphi$ that let me calculate the mean of $\varphi$ for the average degree of progressivity over the years. In section 8 I recalculate the mean of $\varphi$ under the assumption that $\tau_t$ is zero, which then in turn enables me to calculate the mean of $\varphi$ at all different $\tau_t$. I show the results of the estimation in table 9.

<table>
<thead>
<tr>
<th>$\mu_\varphi$</th>
<th>$\sigma_\varphi$</th>
<th>$E[\varphi]$</th>
<th>$\bar{\kappa}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.2850</td>
<td>0.7776</td>
<td>1.0338</td>
<td>-0.2405</td>
</tr>
<tr>
<td>(0.0148)</td>
<td>(0.008)</td>
<td>(0.0099)</td>
<td>(0.0068)</td>
</tr>
</tbody>
</table>

\(^{25}\)Further, I estimate the variances of the innovations to the taste-shifter $b$ that moderates the disutility from work. Again, an initial condition needs to be chosen: I assume that the zeroth and first innovation variance are of the same magnitude. I display these in Appendix 10.7.

The implied estimate for the average Marshallian elasticity across the years can be calculated from the following formula implied by (21) and (22),

$$
\bar{\kappa} = \frac{1 - \bar{\tau}}{\gamma + \bar{\tau}} \left( 1 - E[\varphi] \left( 1 + \frac{1 - \bar{\tau}}{\gamma + \bar{\tau}} \right) \right). \tag{24}
$$

The estimate in table 9 is negative and statistically significant with a point estimate of -0.2405. This estimate is larger in absolute terms than the one shown in Blundell et al. (2016a); that being -0.08. However, their 95-percent confidence band does overlap with mine and their result is most likely driven by the more comprehensive model with a second earner and non-separable preferences. This follows, because I get a very similar result for the Marshallian elasticity compared to mine if I plug their baseline estimates into my formula for the Marshallian elasticity (about -0.27). Accordingly, the difference between the two estimates is not much of a surprise, as some of the neglected issues, like correlated shocks between primary and secondary earner, are bound to be picked up by my estimate.
8 Insurance of Earnings Risk

8.1 Calculating Pass-Through

In the following I provide a calculation of the amount of insurance offered by progressive taxation against the risk stemming from the stochastic process underlying wages. In particular, I quantify how much of a given permanent wage shock transfers onto hours and subsequently earnings.\textsuperscript{26} The following decomposition of the pass-through coefficient of a permanent shock to earnings is:

\[
\frac{\partial \Delta \ln y_t}{\partial \zeta} \approx (1 - \tau_t) \left( 1 + \frac{\partial \Delta \ln h_t}{\partial \zeta} \right)
\]  

Using the structural equations (4),(21) and (22) this expands to,

\[
\frac{\partial \Delta \ln y_t}{\partial \zeta} \approx \left( 1 - \tau_t \right) \left( 1 + \frac{1 - \tau_t}{\gamma + \tau_t} \left( 1 - E[\varphi] \left( 1 + \frac{1 - \tau_t}{\gamma + \tau_t} \right) \right) \right).
\]  

Earnings react with one plus the Marshall elasticity to a given shock and the total response is dampened by the progressivity parameter. If the tax and transfer system did not feature progressivity, the impact on earnings would be

\[
\frac{\partial \Delta \ln y_t}{\partial \zeta} \bigg|_{\tau=0} \approx \left( 1 + \frac{1}{\gamma} \left( 1 - E[\varphi] \left( 1 + \frac{1}{\gamma} \right) \right) \right).
\]  

\textsuperscript{26}Authors often quantify the consumption response as well. However, since I don’t use consumption data, this is not possible. The consumption response is

\[
\frac{\partial \Delta \ln c_t}{\partial \zeta} \approx \frac{\partial \Delta \ln y_t}{\partial \zeta} - \frac{\partial \Delta s/y}{\partial \zeta},
\]

where \( \frac{\partial \Delta s/y}{\partial \zeta} \) is the savings response and \( s/y \) is the average propensity to save.

\textsuperscript{27}The conversion is straightforward.

\[
E[\varphi] = \frac{1 - \pi}{1/\varphi + (1 - \pi)\frac{1 - \pi}{\gamma + \tau}},
\]  

where \( \pi \) is the mean of the ratio of assets to total wealth, which is the sum of assets and human wealth. Then

\[
E \left[ \varphi \right]_{\tau=0} = 1/ \left( 1/\varphi + \frac{(1 + \gamma)\pi}{\gamma(\gamma + \tau)} \right).
\]  

27
the average progressivity parameter and the average standard deviation of the permanent shock.

Table 10. Average Insurance of Earnings

\[
\begin{array}{cccc}
1 - \bar{\tau} & \bar{\sigma}_\zeta & \frac{\partial \Delta \ln y_t}{\partial \zeta} \bar{\sigma}_\zeta & \frac{\partial \Delta \ln y_t}{\partial \zeta} \bigg|_{\tau=0} \bar{\sigma}_\zeta & \%\text{-reduction} \\
0.925 & 0.22 & 0.1538 & 0.1631 & 5.7 \\
\end{array}
\]

Note: Own calculation based on PSID (2015).

A shock of size 0.22 is attenuated by roughly 30 percent, through both the tax system and the labor supply reaction because the Marshallian is negative. When the tax system offers no insurance, the attenuation is only about 25 percent. Therefore, insurance offered through progressive taxation is roughly 5.7 percent. When I shut down the labor supply reaction (\(\frac{\partial \Delta \ln h_t}{\partial \zeta} = 0\)), progressive taxation is the only source of insurance, and the percentage reduction of the shock equals \(\bar{\tau}\), so 7.5 percent. The degree of insurance offered by progressive taxation is attenuated by the labor supply reaction.

I can now calculate the year-specific impact of a permanent wage shock \(\frac{\partial \Delta \ln y_t}{\partial \zeta} \bigg|_{\tau=0}\) and the year-specific degree of insurance. I show these values in table 11.

Table 11. Earnings Pass-Through Values and Insurance by Year

\[
\begin{array}{cccccccc}
\frac{\partial \Delta \ln y_t}{\partial \zeta} \bigg|_{\tau=0} & 2000 & 2002 & 2004 & 2006 & 2008 & 2010 & 2012 \\
0.6820 & 0.6912 & 0.7113 & 0.6962 & 0.7037 & 0.6924 & 0.7054 \\
\sigma_{\zeta,t} & 0.1853 & 0.2011 & 0.1970 & 0.2194 & 0.2503 & 0.2503 & 0.2374 \\
\frac{\partial \Delta \ln y_t}{\partial \zeta} \bigg|_{\tau=0} \sigma_{\zeta,t} & 0.1264 & 0.1390 & 0.1401 & 0.1527 & 0.1761 & 0.1733 & 0.1675 \\
\%\text{-reduction} & 7.3 & 6.0 & 3.3 & 5.4 & 4.4 & 5.9 & 4.1 \\
\end{array}
\]

Note: Own calculation based on PSID (2015). The last line represents the additional insurance above the level of \(\frac{\partial \Delta \ln y_t}{\partial \zeta} \bigg|_{\tau=0}\).

Table 11 relates two important facts. First, the amount of pass-through from permanent wage shocks to income has grown over time and therefore the amount of insurance offered by the tax system has waned. However, the change in the pass-through coefficient is relatively small. The highest pass-through is recorded in 2004 and the lowest in 2000, while overall growth from 2000 to 2012 was about 3 percent. This leads me to conclude that the pass-through coefficient has been rather stable over time even through the crisis. Second, and in stark contrast to the time series of the pass-through coefficient: risk, measures by the standard deviation of permanent wage shocks, has grown quite substantially. \(\sigma_{\zeta,t}\) started at a level of 0.18, grew to a peak of 0.25 during the crisis and fell to 0.23 in 2012; total growth from 2000 to 2012 is about 28 percent. Since the pass-through parameter remained roughly constant, shocks passed on to earnings follow the trend in wage shocks. Pass-through adjusted risk grew from 0.12 in 2000 to 0.16 in 2012, a 33% increase. In sum,
permanent earnings risk rose quite drastically, but the tax system, as captured by \( \tau \), did not undergo major alterations after that.

To characterize the relationship between \( \tau \) and the pass-through coefficient to earnings, I display figure 5.

![Figure 5. Earnings Pass-Through and Marshallian Elasticity as Functions of \( \tau \)](image)

(a) Earnings Pass-Through

(b) Marshallian Elasticity

Figure 5. Earnings Pass-Through and Marshallian Elasticity as Functions of \( \tau \)

*Note*: Own calculation based on PSID (2015). Shows the pass-through coefficient \( \frac{\Delta \ln y}{\Delta \zeta} \) at different levels of \( \tau \) and the pass-through if there were no labor supply response, i.e. \( 1 - \tau \). Also shows the and the Marshallian elasticity \( \kappa \) as a function of \( \tau \).

The upper panel compares the pass-through to earnings as calculated above and the pass-through when the labor supply reaction is set to zero, while the lower panel characterizes the labor supply reaction by graphing the Marshallian elasticity over the range of \( \tau \). At the origin \( \tau = 0 \), so that the pure labor supply reaction can be seen on the abscissa. As \( \tau \) and therefore progressivity increase, pass-through is diminished and insurance increases. The initial rise is slow, as the labor supply reaction runs counter to the new insurance offered through the tax system. However, at very high levels of progressivity, the labor supply reaction becomes less and less important, as the margin to respond becomes thinner and thinner. This leads to the convergence of the two curves at \( \tau = 1 \). This can also be verified in the lower panel. The labor supply reaction, i.e. the Marshallian elasticity, increases with \( \tau \), so that it becomes more muted. Finally the reaction is zero at \( \tau = 1 \).
### 8.2 Stabilizing Earnings Risk

I have shown the rise in permanent earnings risk over the early 2000s and that progressive taxation in terms of the progressivity parameter \(1 - \tau\) played a minor role in shaping how this risk transferred onto earnings. Rather, I find that the labor supply response was of primary importance for the pass-through.

The following is an illustration of the counterfactual approach to tax policy evaluation using the methods of section 6.1.3. I calculate the levels of the progressivity parameter \(1 - \tau\) and the adjustment to the top tax rate in each year that would have resulted in holding the level of earnings risk (\(\frac{\partial \Delta \ln y_t}{\partial \zeta_t} \bigg|_{\tau_t, \sigma_{\zeta,t}}\)) constant at the average level over the sample period, which is about 0.1536.\(^{28}\) The top tax rate is paid by a small fraction of tax units in the dataset, ranging from 655 to 1990 cases (frequency weights applied) in the sample period. This exercise is supposed to illustrate the influence of the policy maker in shaping the risk experienced by individuals. I show the calculations in table 12.

| Year | \(\frac{\partial \Delta \ln y_t}{\partial \zeta_t} \bigg|_{\tau_t, \sigma_{\zeta,t}}\) | \(1 - \tau_t\) avg. risk | \% change of top tax rate to reach \(\tau_t\) avg. risk |
|------|-----------------|----------------|------------------|
| 2000 | 0.1264          | 1.1661         | -47.81           |
| 2002 | 0.1390          | 1.0412         | -29.39           |
| 2004 | 0.1401          | 1.0709         | -45.6            |
| 2006 | 0.1527          | 0.9263         | -1.84            |
| 2008 | 0.1761          | 0.7808         | 56.38            |
| 2010 | 0.1733          | 0.7808         | 38.02            |
| 2012 | 0.1675          | 0.8356         | 36.91            |

Note: Own calculation based on PSID (2015). Last line calculated using the year-specific percentage changes that a 1% change in the top tax rate induces.

At an average passed-on earnings risk of 0.1536, the period most closely resembling this level of risk is 2006, which – necessarily – is the period with the smallest implied change to the top tax rate. All previous periods had lower values of risk and therefore, to achieve stabilization, progressivity has to be decreased by cutting rates quite substantially. In contrast, the periods after 2006 imply changes toward higher progressivity than is observed. The implied progressivity parameter in this period hovers around 0.8, which comes with a strong increase in the top tax rate. However, all the implied changes are to be taken with a grain of salt, as I have calculated only what would be implied for the top tax rate holding all other aspects of the tax system fixed. Certainly, decreases of the top tax rate would be tied to lowering some or all the other rates as well, if a certain change resulted in the top rate falling below the second-highest or other lower rates. However, I cannot take this into account.\(^{29}\)

However, the broad picture in considering the changes around the crisis is clear. The financial crisis of 2008 was accompanied by a substantial rise in permanent wage and

---

\(^{28}\)Note that this exercise is not budget neutral

\(^{29}\)However, as the parameter \(\tau\) shapes the retention function globally, the new system after the policy change does generally imply a change for individuals other than those paying the top tax rate. So in this sense the new approximated system corresponds to a certain tax schedule with lowered tax liabilities for tax units at the lower end of the income distribution.
earnings risk. To mitigate this risk and keep it at the level of 2006, the state would have had to increase progressivity drastically during the crisis. For example, the relative difference between the $1 - \tau_{2008}^\text{risk}$ and $1 - \tau_{\text{avg. risk}}^\text{2008}$ is roughly -16.3%. However, the observed pattern of progressivity during the crisis is the exact opposite. Progressivity generally decreased or at least stayed above the level of 2006, implying that the state did not substantially react to this rise in the riskiness of earnings by altering the tax system.

9 Conclusion

In this paper I document rising permanent earnings risk from 2000 to 2012 in a model of life-cycle labor supply. The increase of the permanent earnings risk is steady over the first half of this period and punctuated by the crisis in 2008. Namely, I document a 14% increase of the standard deviation of the permanent risk component in 2008 compared to the previous period. Further, I shed light on the role that progressive taxation played in the mitigation of this risk, which is minor.

The tax and transfer system is approximated by way of a power function to make the labor supply estimation tractable. Deductions have a minor role in shaping the progressivity of the system, but they do make it less progressive. An intriguing finding is that the relevant parameters of the approximation, especially the parameter $\tau$, which determines progressivity, are sensitive to the estimation method. Estimating the approximation using nonlinear least-squares implies smaller values of $1 - \tau$; 0.93 instead of 0.82 found with the log specification. The fit of the model estimated using nonlinear least-squares performs about two times better in terms of root mean square error compared to the log specification. In general, the power-function fits the data quite well, with an implied tax liability that is on average 650$ higher than the values derived from the tax-simulation model taxsim.

The estimation of the life-cycle labor supply model mostly confirms findings in the related literature. I find a tax-modified Frisch-elasticity of labor supply of 0.469 and an unmodified, after-tax Frisch-elasticity of 0.528. These values locate in the middle of the estimates presented in Blundell et al. (2016a) and Heathcote et al. (2014). The Marshallian elasticity of labor-supply is negative and larger in absolute value compared to Blundell et al. (2016a).

Finally, I find that the pass-through of permanent wage-risk to earnings is roughly constant over time, which implies that earnings risk is mainly driven by the rise of permanent wage risk. From a counterfactual calculation I determine that to drive down earnings risk to pre-crisis levels, the progressivity parameter should have been lowered to 0.78 instead of the observed 0.93. This could have been achieved by raising the top tax rate by 56 percent.

In sum, I find that permanent earnings risk has been increasing at a steady clip over the early 2000s and taken a significant jump after the crisis of 2008 hit. The government, however, has not exercised much influence over this rise and has hardly varied the progressivity of the tax and transfer system.


10 Appendix

10.1 Measurement Error in Hours, Wages and Earnings

Following Blundell et al. (2016a), I correct the measurement error in log earnings, hours and wages using the estimates from the validation study Bound et al. (1994) to determine the proportion of the overall variance that is due to measurement error. Let \( \tilde{y}, \tilde{h} \) and \( \tilde{w} \) denote observed log earnings, hours and wages respectively. Then for any of these the following relationship holds,

\[
\tilde{x} = x + me^x, x \in \{ y, h, w \},
\]

where \( me^x \) denotes the measurement error and \( x \) the true value. From Blundell et al. (2016a) I adopt the following relationships,

\[
\begin{align*}
Var(me^y) &= 0.04 Var(\tilde{y}), \\
Var(me^h) &= 0.23 Var(\tilde{h}), \\
Var(me^w) &= 0.13 Var(\tilde{w}).
\end{align*}
\]

It follows that the covariance between measurement error in wages and hours is given by

\[
\text{Cov}(me^w, me^h) = \frac{1}{2} \left( Var(me^y) - Var(me^h) - Var(me^w) \right)
\]  

With Differenced Variables  The estimation of the stochastic processes is defined in terms of differenced variables, hence I need to account for the differenced and not the contemporary measurement error. In analogue to the above definitions, let \( \Delta \tilde{x} = \tilde{x}_t - \tilde{x}_{t-1} \). Thus,

\[
\Delta \tilde{x} = \Delta x + \Delta me^x.
\]

The variance of the measurement error of differenced earnings is thus given by,

\[
\begin{align*}
Var(\Delta me^y) &= Var(me^y_t) - 2 \text{Cov}(me^y_t, me^y_{t-1}) + Var(me^y_{t-1}) \\
Var(\Delta me^h) &= Var(me^h_t) + Var(me^h_{t-1})
\end{align*}
\]

(31)  (32)

where the second line follows from the assumption that measurement error is not correlated over time. Since the information about measurement error variances available from the validation study only covers relationships in levels, this assumption is required for the correction to be generalizable to temporal differences.

Again, I need to know the covariance of measurement errors in hours and wages to proceed with the estimation. By directly evaluating the covariance I find that,

\[
\text{Cov}(\Delta me^w, \Delta me^y) = \text{Cov}(me^w_t, me^h_t) + \text{Cov}(me^w_{t-1}, me^h_{t-1}).
\]

(33)
### 10.2 Sample Statistics by Year

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>40.25</td>
<td>8.86</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2000</td>
<td>40.80</td>
<td>9.16</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2002</td>
<td>41.04</td>
<td>9.51</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2004</td>
<td>41.01</td>
<td>9.89</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2006</td>
<td>41.14</td>
<td>1.35</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2008</td>
<td>41.02</td>
<td>1.77</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2010</td>
<td>40.75</td>
<td>1.68</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2012</td>
<td>40.44</td>
<td>1.75</td>
<td>25</td>
<td>60</td>
</tr>
<tr>
<td>2014</td>
<td>40.15</td>
<td>1.72</td>
<td>25</td>
<td>60</td>
</tr>
</tbody>
</table>

**Note:** Own calculation based on PSID (2015). All statistics are unweighted.

### 10.3 Approximation of the Euler Equation

The Euler equation of consumption is given by

$$\frac{1}{\rho(1+r_t)} \lambda_t = E_t[\lambda_{t+1}]$$  \hspace{1cm} (34)

Expectations are rational, i.e., $\lambda_{t+1} = E_t[\lambda_{t+1}] + \varepsilon_{\lambda_{t+1}}$, where $\varepsilon_{\lambda_{t+1}}$ denotes the mean-zero expectation correction of $E_t[\lambda_{t+1}]$ performed in period $t + 1$. Expectation errors are caused by innovations in the hourly wage residual $\omega_{t+1}$, which, as implied by rational expectations, are uncorrelated with $E_t[\lambda_{t+1}]$. Rational expectations imply that $\varepsilon_{\lambda_{t+1}}$ is uncorrelated over time, so that regardless of the autocorrelative structure of the shock terms, $\varepsilon_{\lambda_{t+1}}$ will only be correlated with the innovations of the shock processes.

To find an estimable form for $\Delta \ln h_t$, we take logs of (34) and resolve the expectation:

$$\ln \lambda_t = \ln(1 + r_t) + \ln \rho + \ln (\lambda_{t+1} - \varepsilon_{t+1})$$

A first order Taylor-expansion of $\ln (\lambda_{t+1} - \varepsilon_{t+1})$ gives $\ln (\lambda_{t+1}) + \frac{\varepsilon_{\lambda_{t+1}}}{\lambda_{t+1}}$, leading to the expression

$$\ln \lambda_t = \ln(1 + r_t) + \ln \rho + \ln \lambda_{t+1} + \frac{\varepsilon_{\lambda_{t+1}}}{\lambda_{t+1}}$$
\[
\ln \lambda_t = \ln(1 + r_t) + \ln \rho + \ln (\lambda_{t+1}) + \frac{\varepsilon_{t+1}}{\lambda_{t+1}} + O\left(-\frac{1}{2}(\varepsilon_{t+1}/\lambda_{t+1})^2\right).
\] (35)

Accordingly, when we backdate 35, I can remove \(\ln \lambda_t\) in the first difference formulation of the labor supply equation 20.

10.4 A Model with Explicit Expenditure for Deductions

10.4.1 Two-Stage Budgeting

In the following I explore the within-period leisure-consumption-allocations of individuals when they have the ability to deduct from their taxable income by making purchases of deductible goods.

In the case of perfect foresight and with intertemporally additively-separable utilities over the life-cycle one can decompose the standard optimization of the consumer into two separate optimization procedures (see Blundell and Walker (1986) and Keane (2011)). First, the consumer allocates full income \(F_t\) into each period out of life-time wealth \(W_t\) so as to equate the appropriately discounted values of the marginal utility of income. Second, the consumer chooses in-period consumption-leisure bundles to maximize in-period utility. I derive in an expression for the optimal choice of hours and whether a change in the wage \(w_t\) influences hours also through the parameters that relate expenditure for deductible goods to deductions.

10.4.2 In-Period Allocation

The optimization of the in-period utility function is the same as in any static problem with the exception that it proceeds with full income given, so that \(F_t\) is fixed. The optimization program is

\[
\max_{c_t, c^d_t, h_t} U(c_t, c^d_t, h_t),
\]

\text{s.t. } \quad F_t = \chi \left( w_t T - \mathcal{D}\left(p_{d,t} c^d_t\right)\right)^{1-\tau} - \chi \left( w_t h_t - \mathcal{D}\left(p_{d,t} c^d_t\right)\right)^{1-\tau} + c_t + p_{d,t} c^d_t
\] (37)

where, for simplicity’s sake, I use the parametric form

\[
\mathcal{U}(c_t, c^d_t, h_t) = u \left( c^d_t, c_t \right) - \frac{h_t^{1+\gamma}}{1+\gamma}, \quad \gamma \geq 0.
\] (38)

Here I divide consumption into two types of goods, non-tax-deductible goods \(c_t\), with its price normalized to 1, and tax-deductible goods \(c^d_t\) and price \(p_{d,t}\). Utility from consumption is given by the concave and twice continuously differentiable function \(u(\cdot)\). The function \(\mathcal{D}(\cdot)\) gives the deductions from gross income and is increasing in \(p_{d,t} c^d_t\). Further, \(T\) is the total time endowment in the period. To find an indication of the role that deductible goods play in determining labor supply, I have to inspect the first-order condition for \(h\):
\[ h_t^\gamma = \lambda \left( \chi (1 - \tau) \left( w_t h_t - D \left( p_{d,t} c_{d,t}^d \right) \right)^{-\tau} \right) w_t. \] 

(39)

This expression reveals that, the choice of hours depends nonlinearly on deductions.

Building on this result, I make the pragmatic choice in section 4.3 of approximating taxable income as a power function of gross income. Otherwise there would be no way of proceeding with the first-order approach or the impact of deductions would have be neglected.

10.5 Replication Estimation of the Tax Function Approximation in Heathcote et al. (2017a)

I estimate the model for statutory progressivity shown in eq. A2 of Heathcote et al. (2017b) using the data provided in the replication files for Heathcote et al. (2017a). This means, I model their net income variable minus deductions based on taxable income. I do it once in logs and once using nonlinear least-squares with the quantities in levels. In line with their estimation procedure I pool all observations in their panel from 2000 to 2006. Each cross-section contains between 3000 to 3500 observations. The results are shown in table 14. I use their notation, where \( 1 - \tau \) is the progressivity parameter and \( \lambda \) is the coefficient of the function equivalent to \( \chi \) in my notation.

<table>
<thead>
<tr>
<th></th>
<th>log spec.</th>
<th>nonlin. spec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda )</td>
<td>5.568</td>
<td>1.449</td>
</tr>
<tr>
<td>( 1 - \tau )</td>
<td>0.819</td>
<td>0.937</td>
</tr>
<tr>
<td>RMSE</td>
<td>18714</td>
<td>8932</td>
</tr>
<tr>
<td>Obs.</td>
<td>12875</td>
<td>12875</td>
</tr>
</tbody>
</table>

The table shows that the nonlinear specification and the log specification imply very different progressivity estimates. The nonlinear model implies much lower progressivity. Further, the fit, assessed by computing the RMSE of the predicted residuals in levels, is slightly more than twice as large when computed based on the log specification.

In table 15 I show the estimates year by year along with my own estimates for the total progressivity. The take-away is that the estimates of the statutory progressivity parameter, when measured with nonlinear least-squares, are very close to the estimates of total progressivity that I calculate. Hence, I find that the quantitative importance of the distinction between statutory and total progressivity is quite small. However, the table does highlight the inferior fit of the log specification in terms of the RMSE. In every year the log specification fits worse than the equivalent nonlinear specification.

10.6 Identification Example of the Stochastic Process for Wages

Let the stochastic process for wages be
The equations giving the value of the time-differenced innovation \( \Delta \omega_{it} \) are

\[
\Delta \omega_{it} = \begin{cases} 
(\theta-1)\epsilon_{it-1}+\zeta_{it}+\epsilon_{it}+\Delta m_{ei}t & \text{if } t=1 \\
(\theta-1)\epsilon_{it-1}-\theta\epsilon_{it-2}+\zeta_{it}+\epsilon_{it}+\Delta m_{ei}t & \text{if } t>1 
\end{cases}
\]

(41)

Let there be three periods, such that \( t \in \{1, 2, 3\} \). Then I obtain the following matrix of autocovariance moments:

\[
VCC_{\Delta \omega} = \\
\begin{pmatrix} 
2\sigma_{m2}^2 + (\theta-1)^2\sigma_{e0}^2 + \sigma_{e1}^2 + \sigma_{e2}^2 & -\sigma_{m2}^2 - \theta^2 \sigma_{e0}^2 - \sigma_{e1}^2 + \theta(\sigma_{e0}^2 + \sigma_{e1}^2) & -\theta \sigma_{e1}^2 \\
-\sigma_{m2}^2 - \theta^2 \sigma_{e0}^2 - \sigma_{e1}^2 + \theta(\sigma_{e0}^2 + \sigma_{e1}^2) & 2\sigma_{m2}^2 + \sigma_{e1}^2 - 2\theta \sigma_{e1}^2 + \theta^2(\sigma_{e0}^2 + \sigma_{e1}^2) + \sigma_{e2}^2 + \sigma_{e3}^2 & -\sigma_{m2}^2 - \theta^2 \sigma_{e0}^2 - \sigma_{e1}^2 + \theta(\sigma_{e0}^2 + \sigma_{e1}^2) + \sigma_{e2}^2 + \sigma_{e3}^2 \\
-\sigma_{m2}^2 - \theta^2 \sigma_{e0}^2 - \sigma_{e1}^2 + \theta(\sigma_{e0}^2 + \sigma_{e1}^2) & -\sigma_{m2}^2 - \theta^2 \sigma_{e0}^2 - \sigma_{e1}^2 + \theta(\sigma_{e0}^2 + \sigma_{e1}^2) + \sigma_{e2}^2 + \sigma_{e3}^2 & 2\sigma_{m2}^2 + \sigma_{e2}^2 - 2\theta \sigma_{e2}^2 + \theta^2(\sigma_{e1}^2 + \sigma_{e2}^2) + \sigma_{e3}^2 + \sigma_{e4}^2 
\end{pmatrix}
\]

(42)

There are six unique moments in the above matrix that are used for identification. Let every element of \( VCC_{\Delta \omega} \) be denoted by the symbols \( \Gamma_{k,j} \), where \( k,j \in \{1, 2, 3\} \), so that for example the variance in period 1 is \( \Gamma_{1,1} = 2\sigma_{m2}^2 + (\theta-1)^2\sigma_{e0}^2 + \sigma_{e1}^2 + \sigma_{e2}^2 \). The easiest way to proceed is to set the innovation variance of the transitory process in \( t = 0 \) to zero. Further, as laid out in appendix 10.1, I can treat the variance of the measurement error as known. Then the set of moments used for identification becomes:

Table 15. Progressivity Estimates by Year

<table>
<thead>
<tr>
<th>Year</th>
<th>log specification</th>
<th>nonlinear specification</th>
<th>complete model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \lambda )</td>
<td>( 1 - \tau )</td>
<td>RMSE</td>
</tr>
<tr>
<td>2000</td>
<td>4.284</td>
<td>0.840</td>
<td>7229</td>
</tr>
<tr>
<td>2002</td>
<td>6.05</td>
<td>0.811</td>
<td>12953</td>
</tr>
<tr>
<td>2004</td>
<td>6.221</td>
<td>0.811</td>
<td>32099</td>
</tr>
<tr>
<td>2006</td>
<td>5.789</td>
<td>0.818</td>
<td>13402</td>
</tr>
</tbody>
</table>

Note: The first two models are the same as the two displayed in table 14, but are disaggregated by year. The last model is the complete model of table 3 displayed for comparison.
\( \Gamma_{1,1} = \sigma_{\epsilon,1}^2 + \sigma_{\zeta,1}^2 \)
\( \Gamma_{1,2} = -\sigma_{\epsilon,1}^2 + \theta \sigma_{\zeta,1}^2 \)
\( \Gamma_{1,3} = -\theta \sigma_{\epsilon,1}^2 \)
\( \Gamma_{2,2} = \sigma_{\epsilon,1}^2 - 2\theta \sigma_{\epsilon,1}^2 + \theta^2 \sigma_{\epsilon,1}^2 + \sigma_{\epsilon,2}^2 + \sigma_{\zeta,2}^2 \)
\( \Gamma_{2,3} = -\theta^2 \sigma_{\epsilon,1}^2 - \sigma_{\epsilon,2}^2 + \theta (\sigma_{\epsilon,1}^2 + \sigma_{\epsilon,2}^2) \)
\( \Gamma_{3,3} = \sigma_{\epsilon,2}^2 - 2\theta \sigma_{\epsilon,2}^2 + \theta^2 (\sigma_{\epsilon,1}^2 + \sigma_{\epsilon,2}^2) + \sigma_{\zeta,3}^2 + \sigma_{\zeta,3}^2 \)

Accordingly, the identification proceeds by solving for the variances and the persistence parameter

\( \sigma_{\epsilon,1}^2 = -(\Gamma_{1,2} + \Gamma_{1,3}) \)
\( \theta = \Gamma_{1,3}/(\Gamma_{1,2} + \Gamma_{1,3}) \)
\( \sigma_{\zeta,1}^2 = \Gamma_{1,1} + (\Gamma_{1,2} + \Gamma_{1,3}) \)
\( \sigma_{\epsilon,2}^2 = \frac{1}{\Gamma_{1,3} / \Gamma_{1,2} + \Gamma_{1,3} - 1} \left[ \Gamma_{2,3} + \left( \frac{\Gamma_{1,3}}{\Gamma_{1,2} + \Gamma_{1,3}} - \frac{\Gamma_{1,3}}{\Gamma_{1,2} + \Gamma_{1,3}} \right) \times (-\Gamma_{1,2} + \Gamma_{1,3}) \right] \)
\( \sigma_{\zeta,2}^2 = \frac{1}{\Gamma_{1,3} / \Gamma_{1,2} + \Gamma_{1,3} - 1} \left[ \Gamma_{2,3} + \left( \frac{\Gamma_{1,3}}{\Gamma_{1,2} + \Gamma_{1,3}} - \frac{\Gamma_{1,3}}{\Gamma_{1,2} + \Gamma_{1,3}} \right) \times (-\Gamma_{1,2} + \Gamma_{1,3}) \right] \)

Each further period delivers two more moments that can be used for identification, so that the next set of permanent and transitory variances can be identified.

Another possibility to identify the process is to set the first two transitory shock variances equal to each other, \( \sigma_{\epsilon,0}^2 = \sigma_{\epsilon,1}^2 \). The identification proceeds analogously, except that instead of \( \theta \) being identified directly, the ratio \( \theta / (1 - \theta) \) is identified. However, this makes no difference in practice.

### 10.7 Innovations to Taste-Shifter \( b \)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{\nu,t} )</td>
<td>0.3188</td>
<td>0.0370</td>
<td>0.0009</td>
<td>0.4594</td>
<td>0.4237</td>
<td>0.3263</td>
<td>0.5023</td>
</tr>
<tr>
<td>(0.0033)</td>
<td>(0.0044)</td>
<td>(0.0006)</td>
<td>(0.0025)</td>
<td>(0.0027)</td>
<td>(0.0032)</td>
<td>(0.0021)</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Own calculation based on PSID (2015). Bootstrap standard errors based on 200 replications in parentheses.*

The standard deviations for the innovations of the taste-shifter are roughly two times as large as the permanent shock variances of the wage process. There is no very clear-cut trend across the years. In 2004 the estimate is so small that it turns insignificant. This
points to a large degree of instability in the evolution of the variance of the taste-shifter. In the current model this component of the hours variance is purely transitory.

References

Altonji, J. G. and L. M. Segal

Attanasio, O. and J.-V. Rios-Rull

Attanasio, O. P. and G. Weber

Blundell, R., M. Graber, and M. Mogstad

Blundell, R. and T. MacCurdy

Blundell, R., L. Pistaferri, and I. Preston

Blundell, R., L. Pistaferri, and I. Saporta-Eksten

Blundell, R., L. Pistaferri, and I. Saporta-Eksten
2016b. Consumption inequality and family labor supply: Online appendix.

Blundell, R. and I. Walker

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers

Domeij, D. and M. Flodén

Eaton, J. and H. S. Rosen

Feenberg, D. and E. Coutts

Feldstein, M. S.

Gouveia, M. and R. P. Strauss

Heathcote, J., K. Storesletten, and G. L. Violante
Heathcote, J., K. Storesletten, and G. L. Violante

Heathcote, J., K. Storesletten, and G. L. Violante

Jakobsson, U.

Jessen, R. and J. König
2018. Hours risk, wage risk, and life-cycle labor supply. mimeo.

Kaplan, G.

Keane, M. P.

Kimberlin, S., H. L. Shaefer, and J. Kim

Kleibergen, F. and R. Paap

Krueger, D.

Low, H. and D. Maldoom

PSID

PSID

Varian, H. R.