Hours and Wages*

Alexander Bick
Arizona State University

Adam Blandin
Virginia Commonwealth University

Richard Rogerson
Princeton University

First Draft: November 2018, Current Draft: September 2021

Abstract

We document two robust features of the cross-sectional distribution of usual weekly hours and hourly wages. First, usual weekly hours are heavily concentrated around 40 hours, while at the same time a substantial share of total hours come from individuals who work more than 50 hours. Second, mean hourly wages are non-monotonic across the usual hours distribution, with a peak at 50 hours. We develop and estimate a model of labor supply to account for these features. The novel feature of our model is that earnings are non-linear in hours, with the extent of nonlinearity varying over the hours distribution. Our estimates imply significant wage penalties for individuals that deviate from 40 hours in either direction, leading to a large mass of individuals that work 40 hours and are not very responsive to shocks. This has important implications for the role of labor supply as a mechanism for self-insurance in a standard heterogeneous agent-incomplete markets model and for empirical strategies designed to estimate labor supply parameters.

*We thank the editors and three anonymous referees, seminar/conference participants at ASU, the Richmond Fed, Atlanta Fed, St. Louis Fed, Minneapolis Fed, McGill University, Goethe University, University of Delaware, the Bank of Canada, UCL, EIEF, USC, NYU Stern, Chicago, Carlos III, University of Mannheim, LMU, William and Mary, the 2018 SF Fed Micro Macro Labor Conference, CASEE Conference on Labor Markets, the 2019 Southern Economics Association, the 2019 SED, and the 2020 Barcelona GSE Summer Forum for helpful comments and suggestions. We thank Diego Daruich, Gary Hansen, John Kennan, Derek Neal, Josep Pijoan-Mas and Aysegul Sahin, for comments on an earlier version. We would like to thank Spencer Perry, Siyu Shi, and especially Minju Jeong for outstanding research assistance.

1 Introduction

Recent work in macroeconomics emphasizes the desirability of deriving aggregate implications from models that also capture the salient aspects of cross-sectional heterogeneity found in the data. As Krueger et al (2010) wrote in their introduction to the special issue of the Review of Economic Dynamics devoted to this topic, "…restricting heterogeneous agent macro models so that the equilibrium distributions of hours worked, income, consumption and wealth line up well with their empirical counterparts is crucial for a convincing policy analysis." But what set of statistics represents the key cross-sectional facts? We pursue this question in the context of heterogeneous agent models of labor supply. Central to any study of labor supply is the relation between wages and hours. While the existing literature on heterogeneous agent models of labor supply has tended to focus on first and second moments of the cross-sectional hours and wage distribution (see, for example, Heathcote et al (2014)), we focus on two features of the micro data not captured by these moments, and argue that addressing them has important implications not only for the role of labor supply in heterogeneous agent macro model but also for standard empirical strategies to estimate labor supply parameters.

The first feature is that usual weekly hours are heavily concentrated around 40 hours, while at the same time a significant share of workers have usual hours of 50 or more. The second feature is the non-monotonic relationship between wages and usual weekly hours: mean wages increase until 50 hours and decrease after. In particular, we show that hourly wages for those working 65 hours are about the same as for those working 35 hours, which are about 10 log points below the hourly wage at 40 hours. Our emphasis on this cross-sectional relationship is a novel feature of our analysis, and one contribution of our paper is to document that the non-monotonic relation between hourly wages and usual weekly hours is a robust feature of the data; it holds when we cut the data by gender, age, education, industry and occupation. It also holds across every major US dataset with data on earnings and weekly hours of work and we argue it is not primarily an artifact of measurement issues.[1]

Textbook static labor supply models assume that earnings are linear in hours worked. Although a version of this model in the spirit of Heathcote et al (2014) can account for the first and second moment properties of the cross-sectional hours-wage distribution, it cannot account for the two features described above. This motivates us to develop an extension of the simple textbook model that allows earnings to depend non-

[1] A large literature on the part-time wage penalty focuses on the increasing profile below 40 hours. Some work examines wages associated with long hours, including, for example, Hirsch (2005), Kuhn and Lozano (2008), Michelacci and Pijoan-Mas (2012), Goldin (2014), Weeden et al (2016), Yurdagul (2017), Gicheva (2020), Denning et al (2019), Fuentes and Leamer (2019) and Shao et al (2021).

linearly on hours. Our specification generalizes the one introduced by French (2005); whereas he assumed an elasticity of earnings with respect to hours that was constant and different from one, we let this elasticity vary with hours of work. That is, our specification can accommodate the possibility that an individual moving from 40 hours to 30 hours incurs a part-time wage penalty without requiring that this same effect is operative for a worker moving from 50 hours to 40 hours.

Our model is able to replicate the key features of the cross-sectional wage and hours data that we document in our empirical analysis. As a starting point, we illustrate this for a sample of men aged 50-54 with a high school degree. We choose males 50-54 because our model abstracts from dynamic considerations such as human capital accumulation, and dynamic considerations are commonly assumed to be less important for workers in this age group. (See, for example, Heckman et al, 1998). Next, we show that varying the extent of the non-linearity but holding all other parameters fixed allows our model to account for the cross-sectional patterns for males aged 50-54 with either college or post-college degrees. Our estimated model also does a good job of accounting for the cross-sectional wage and hours patterns for younger age groups, suggesting that the forces captured by our model estimated on older workers play an important role in shaping outcomes for all workers.

Our estimated model features two key properties. First, selection on unobservables plays an important role in shaping the profile of observed wages across the hours distribution; in our estimated model, individuals working part-time are more likely to be low productivity and individuals working long hours are more likely to be high productivity. Thus, the wage-hours choices faced by an individual worker cannot be inferred directly from the cross-sectional relationship between wages and hours, and requires a structural model of labor supply.

Second, there is a large kink in earnings at 40 hours per week, with the elasticity of earnings with respect to hours dropping from well above one to well below one. These elasticities imply not only a wage penalty for part-time work but also a wage penalty for increasing hours beyond 40. The part-time wage penalty that we estimate is consistent with existing evidence, including, for example, the estimate of Aaronson and French (2004) for older males. Our estimated wage penalty for long hours is substantial: an individual who chooses to work 50 rather than 40 hours faces a wage penalty of almost 20 percent. Our estimate of this (static) long hours penalty is similar in magnitude to the one estimated by Michelacci and Pijoan-Mas (2012) using a dynamic panel analysis of PSID data. The kinked relationship between earnings and hours plays a key role in generating both the concentration of hours at 40 and the non-monotonicity in the wage-hours profile. Moreover, the large mass of individuals who are located at a kink in the earnings function are much less likely to adjust hours worked in response to changes in wages, other income and tax rates. This has important implications for labor supply, which we illustrate in several applications.

In the first application, we embed our estimated kinked relationship between hours and earnings into an otherwise standard heterogeneous agent-incomplete markets model with endogenous labor supply, as studied by Pijoan-Mas (2006). His model assumes a linear relationship between hours and earnings, and he finds that labor supply responses to productivity shocks play a quantitatively large role in helping to smooth consumption. Our kinked specification significantly decreases the potency of this margin, but importantly, this effect is not uniform across the usual hours worked distribution. Individuals who typically work at or near the kink at 40 hours hardly adjust their labor supply in response to temporary variation in productivity, and instead rely solely on asset accumulation for self-insurance. In contrast, workers typically working away from the kink at 40 hours behave more similarly to those in Pijoan-Mas (2006).

The remaining applications all concern empirical strategies designed to estimate preference parameters that shape labor supply responses. The first strategy relates to the large literature that estimates Frisch elasticities using panel data (see, e.g., MaCurdy 1981, Browning et al 1985, and Altonji 1986). Our model suggests that individuals located at the large kink at 40 hours will have smaller responses to temporary wage changes. We repeat the estimation exercises in Bredemeier et al (2019) using data from the PSID, but split the sample based on average hours of work. The estimated Frisch elasticities display a U-shaped pattern with respect to average hours, with the smallest values for those who have average weekly hours that are close to 40. Viewed through the lens of our model, this reflects reduced opportunities for intertemporal substitution and not necessarily a reduced willingness to do so. Because top earners tend to work more than 40 hours, these implications may impact optimal tax policies.

Our results may also affect the interpretation of findings from the literature that relies on bunching at kinks generated by the tax and transfer system to estimate labor supply elasticities, see e.g., Blomquist (2021) for a recent study. Our estimated model generates an order of magnitude less bunching at a kink in the tax schedule relative to a model in which earnings are linear in hours. The reason for this is that most regions within the male earnings distribution are dominated by individuals that work 40 hours, and so are located at a large kink in the earnings-hours relationship. Viewed through the lens of our model, the small amount of bunching at a kink in the tax schedule need not reflect a small labor supply elasticity; rather it may simply reflect a large kink in the earnings-hours relationship relative to kinks generated by the tax system.

Finally, our model has implications for the literature that uses labor supply responses of individuals with lottery winnings as a way to estimate preference parameters that determine income effects (see, e.g., Cesarini et al (2017). We show that income from sources other than a worker's own labor earnings has a much smaller effect on hours in our model than in a standard model with linear earnings. This result suggests that some subtlety is required when interpreting the empirical evidence from lottery winnings; through the lens of our model, the individuals located at the kink are much less likely to adjust their hours because of the shape of the earnings-hours relationship rather than primitive features of preferences. Once again, this effect is non-uniform across the hours distribution, and may have important implications for how government transfers affect hours along the intensive margin.

Our paper relates to several strands of the literature on labor supply, including the large literature on estimating labor supply elasticities touched upon above (see the survey by Keane (2011) for additional references). Rosen (1976) and Moffitt (1984) are early examples of empirical studies incorporating a non-linear earnings function and emphasizing its role for labor supply responses.[2] Their focus was the part-time wage penalty in the context of married female labor supply. Michelacci and Pijoan-Mas (2012) allow current hours to affect both current and future wages, though they focus on the dynamic effects. Yurdagul (2017) documents a hump-shaped pattern for wages across the hours profile and studies a production structure in which workers are complements, implying that wages decrease as hours move away from mean hours. Relative to these studies we seek to account for the joint distribution of hours and wages, and thereby also emphasize the role of selection. We also emphasize implications for labor supply responses and the estimation of labor supply elasticities.

Following Cogan (1981), many researchers have posited fixed costs or other non-convexities as a way to account for the fact that the hours distribution has little mass at low hours and a mass of workers working zero hours. (See, for example, French 2005, Rogerson and Wallenius 2009, Erosa et al 2016, Chang et al 2019 and Ameriks et al 2020.) Although these papers generate a distribution of hours among the employed, none of them generates the large concentration of usual weekly hours around 40, nor do they address the cross-sectional wage-hours profile that is a focal point of our analysis.

Heathcote et al (2014) and Chang et al (2020) study heterogeneous agent macro models that address features of the cross-sectional distribution of wages and hours. But they focus on second moment properties and do not account for either the concentration in the hours worked distribution or the non-monotonicity in the cross-sectional wage-hours profile.

An outline of the paper follows. Section 2 documents the key facts that are the focal point of our analysis. Section 3 presents our model of labor supply that features kinks in the mapping from hours to earnings, and describes the moment matching exercise used to estimate our model. Section 4 presents the main results of our estimation, showing that our model provides a good fit to the data for different age and education groups

[2] See Barzel (1973) and Rosen (1978) regarding the general notion of wages that depend on hours. There is a large literature starting with Hausmann (1985) on econometric estimation of models with nonlinear budget sets.

and that it performs much better than standard alternatives. Section 5 extends our analysis to allow for an additional dimension of heterogeneity, namely other sources of income. Section 6 describes the implications of our model for the responsiveness of hours to changes in wages. Section 7 extends the analysis to allow for compensation not reflected in current earnings and Section 8 concludes.

2 Cross-Sectional Facts About Hours and Wages

The cross-sectional properties of the hours and wage distributions are often summarized by the covariance matrix for log hours and log wages, which amounts to reporting three statistics: the standard deviation of log wages, the standard deviation of log hours and the correlation between the two. In this section we document a broader set of cross-sectional facts. First, the distribution of usual weekly hours across individuals features a heavy concentration around 40, while at the same time a large amount of total hours are accounted for by those who work 50 or more hours per week. Second, we document a striking fact about the profile of average hourly wages across the usual weekly hours distribution: it is non-monotonic with a peak occurring at about 50 hours.

2.1 Data

The facts that we present in this section are derived from pooling the CPS outgoing rotation group (ORG) surveys from September 1995 through August 2007. We pool multiple years to ensure sufficient sample sizes when we stratify the data by various characteristics. We start in 1995 since this allows us to see whether earnings are imputed. We stop in 2007 to avoid the potential concern that our results are impacted by the Great Recession; in fact, the patterns we document continue to hold in later years. As documented below, the patterns that we document also hold quantitatively in many other datasets (CPS ASEC, i.e., the March Supplement of the CPS, the Census, the ACS, the PSID, the NLSY79 and the SCF).[3]

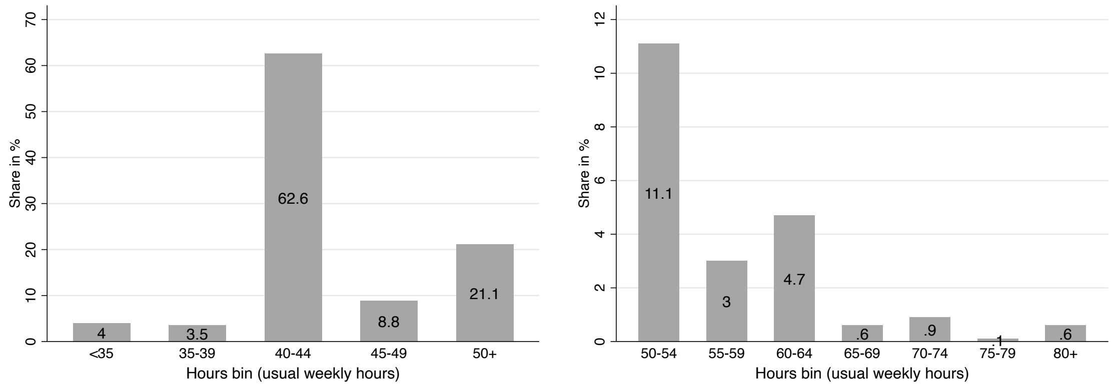
The two key variables from the CPS ORG that we use in our analysis are usual weekly hours and usual weekly earnings for an individual's main job. For the main results presented below we restrict attention to males between the ages of 25 and 64 who hold a single job (roughly 95% of all workers), have usual weekly hours of at least 10, are not enrolled in school and are not self-employed. We eliminate any observations with imputed values for either usual weekly hours or earnings, or that have an implied wage (earnings/hours) less than one half of the federal minimum wage. This leaves us with a sample of more than half a million observations.

[3] We work with the IPUMS version of the CPS, Census, ACS, and ATUS, see Flood et al (2020), Ruggles et al (2021) and Hofferth et al (2020).

Figure 1: The Distribution of Usual Weekly Hours

(a) Overall Distribution

(b) Long Hours Distribution



Source: CPS ORG September 1995-August 2007.

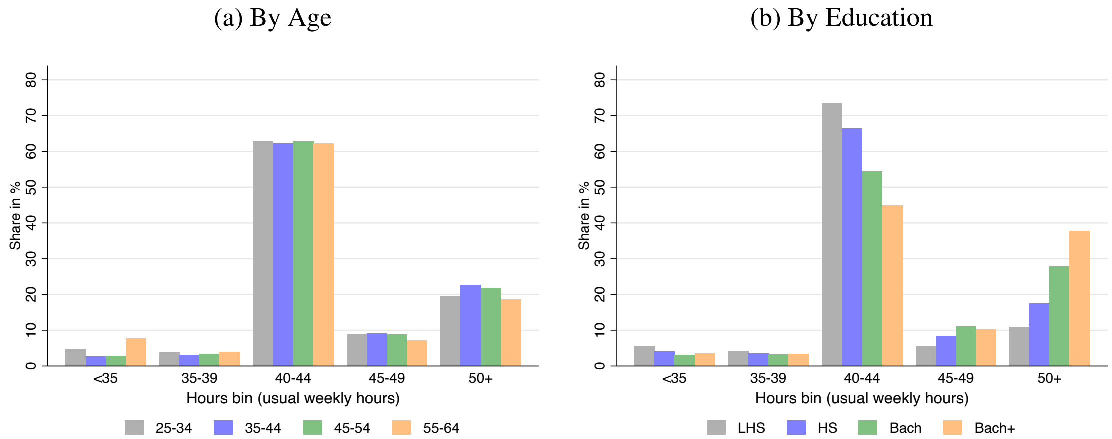
Our focus on males is motivated by two considerations. First, to simplify the analysis, our theoretical model will abstract from the participation margin, and this is much less problematic for males. Second, females are much more likely to work less than 30 hours and less likely to work more than 50 hours. This makes the female population more relevant for studying the part-time wage penalty. Although our analysis includes part-time work, our focus on wages and hours in the long hours region is more novel, thus making the male subsample more relevant. However, while our estimation exercise focuses on males, the interesting empirical patterns we document for males also hold for females. Empirical properties for females are reported in Online Appendix A.

2.2 The Hours Distribution

We start by presenting information on the distribution of usual weekly hours in Figure 1. We note three features. First, there is a heavy concentration in the 40-44 hours bin, with over 60 percent of males reporting usual hours in this range. Almost all of these report usual hours of exactly 40. Second, long hours, which we define as 50 or more, are relatively common, accounting for more than 20 percent of observations, and 28 percent of the total usual hours for men. As the right panel shows, most of these individuals have usual hours below 65. Third, short hours, which we define as less than 35, are relatively uncommon, accounting for only about 4 percent of observations.

Our structural estimation exercise later in the paper will stratify the data by age and education. For this reason we examine how the distribution varies with these two observables in Figure 2. The left panel shows that the three features noted above continue to hold when we stratify by age. The most notable difference across these age groups is that individuals aged 55-64 are somewhat more likely to work short hours and

Figure 2: The Distribution of Usual Weekly Hours By Age and Education



Source: CPS ORG September 1995-August 2007. *Notes:* Our education classification is based on the highest degree obtained. LHS: no high school diploma or equivalent; HS: high school diploma or equivalent, some college but no degree, associate's degree; Bach: Bachelor's degree; Bach+: master's, professional, doctoral degree.

less likely to work 45 or more hours. The right panel of Figure 2 presents the distribution of usual weekly hours by education. In this and subsequent calculations, individuals with some college are included in the high school category. The dominant pattern here is that as educational attainment increases, we essentially move mass from the 40-44 hours bin into the more than 50 hours bin. Nonetheless, it remains true even for the bachelor plus group that there is a concentration in the 40-44 hours bin. And even for the high school dropouts more than ten percent work long hours.

We have also examined the hours distribution by occupation and industry. While the propensity for long and short hours does vary by industry and occupation, it remains true that the largest mass is in the 40-44 hour bin for all two digit occupations and industries.

2.3 Wage-Hours Profiles

In this subsection we study the relationship between wages and hours and how it varies across the hours distribution. We partition the range of weekly hours between 10 and 99 into a set of 5-hour bins: $10-14$, $15-19$, ..., $75-79$, and the final bin running from $80-99$ because there are so few observations there. We denote the set of bins by $H = \{10, 15, ..., 80\}$, where $h \in H$ denotes the minimum threshold of a particular hours bin, and define a set of individual hours dummies $\mathbb{1}_{ih}$ which equal one if individual i's usual weekly hours lies in bin h. We begin our analysis by considering the following regression:

$$w_i = a_0 + \left(\sum_{h \in H} \beta_h \mathbb{1}_{ih} \right) + \gamma X_i + \varepsilon_i \,, \tag{1}$$

where i denotes an individual, and w denotes log hourly wages, defined as the log of usual weekly earnings divided by usual weekly hours. For our first set of results X is a vector of controls including a quadratic in age and dummies for education (less than high school, high school, bachelor's degree, and graduate degree), marital status (married, non-married), race/ethnicity (black, Hispanic, and non-black, non-Hispanic), sector of employment (public and private), union membership, metro area status, state of residence, interview month, and year. Later we will also run the regression separately for various subgroups.

The coefficients of interest are the β_h, which describe how average wages vary across the hours bins. Because so much of our sample falls in the 40 hours bin, we view it as a natural reference point and so omit the β_{40} coefficient from the regression. That is, we normalize β_{40} to be zero, and interpret the remaining β_h as the difference in average log wages in hours bin h relative to the 40 hours bin. We emphasize that we do not attach any causal significance to this estimated relationship – later in the paper we outline a strategy to estimate the underlying causal relationship by using the β_h as moments to be matched in the context of a structural model of labor supply.

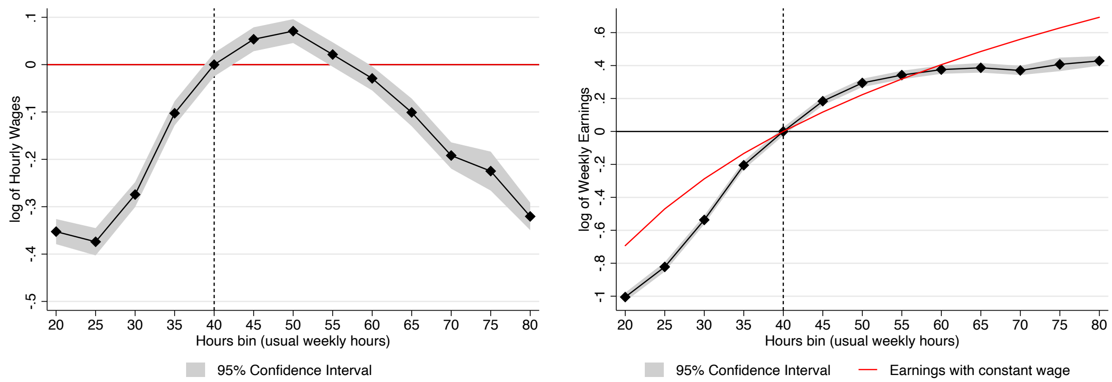
This regression generalizes the analysis in Goldin (2014), Cortes and Pan (2018), and Denning et al (2019). These papers use log weekly hours as a right hand side regressor, thereby implicitly estimating a constant elasticity relationship between the dependent variable (wages in our case and earnings in theirs) and hours. In contrast, our specification allows this elasticity to vary with hours. As we show below, the assumption of a constant elasticity hides important non-linearities in the underlying data. Whether one uses log earnings or log wages as the dependent variable is of no substantive consequence. In the constant elasticity regression, there is a one to one mapping between the implied regression coefficients, and this would also be the case for our specification if we were to use one hour bins. This one-to-one mapping is independent of the well-known division bias coming from measurement error in hours, first discussed by Borjas (1980). Later we provide evidence that measurement error in hours as well as other measurement issues are not the key driver behind our findings.

The left panel of Figure 3 plots the estimates of β_h. As we move from the 25 hours bin to the 40 hours bin, hourly wages increase; hourly wages in the 40 hours bin are more than 20 log points higher than hourly wages in the 30 hours bin. Hourly wages continue to rise, albeit at a slower pace until the 50 hours bin, after which they decrease at a roughly constant rate. Hourly wages in the 55 hours bin are roughly the same as in the 40 hours bin, and hourly wages in the 65 hours bin are about the same as in the 35 hours bin.

The right panel of Figure 3 shows the results when we instead use usual weekly earnings as the left hand side variable. The red line indicates how earnings at each level of hours would compare to the 40 hours bin if there were a unitary elasticity, i.e., if wages were constant across the hours distribution. This figure clearly

Figure 3: Cross-Sectional Relationship between Wages/Earnings and Hours

(a) log Hourly Wages (b) log Weekly Earnings



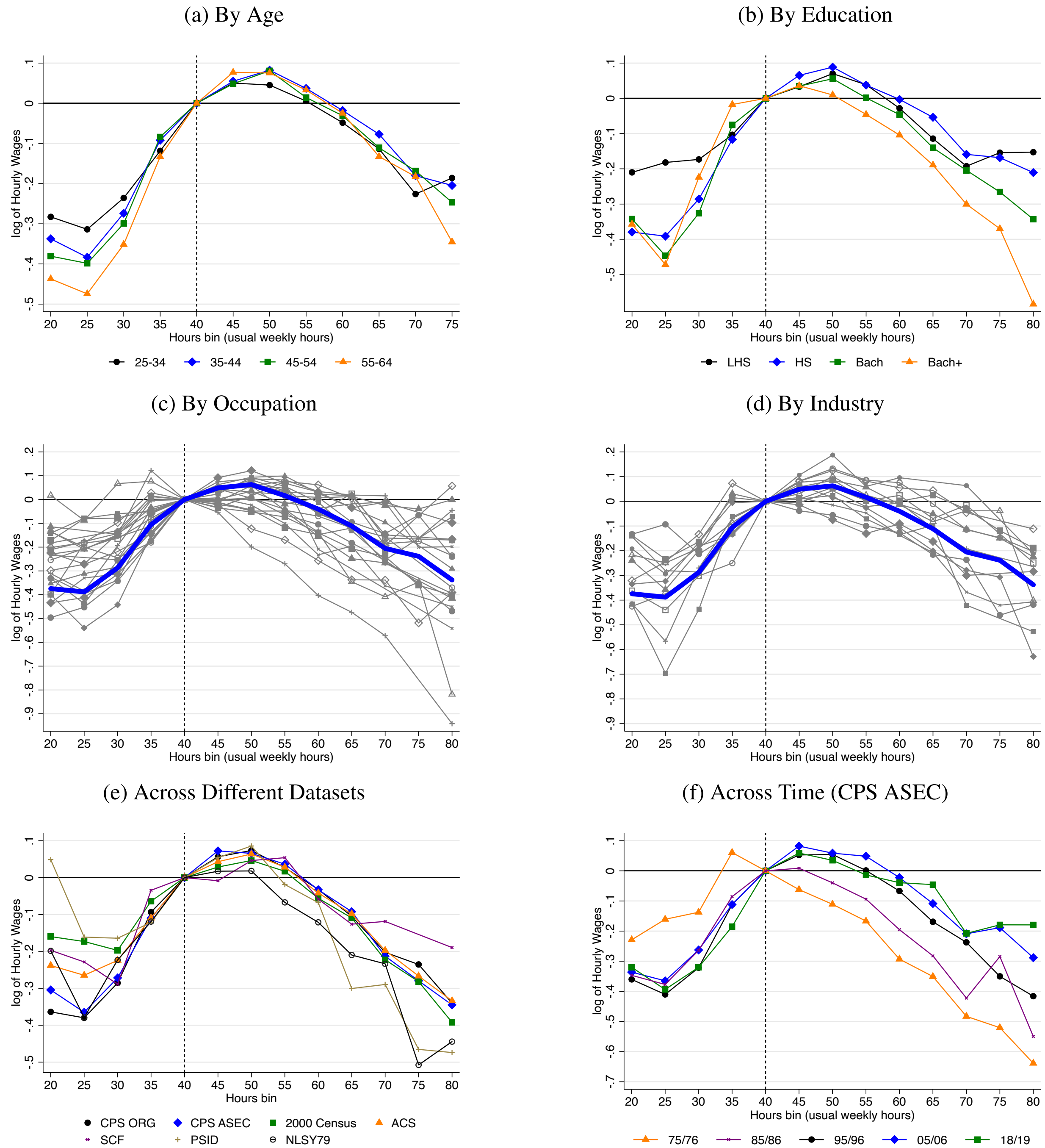
Source: CPS ORG September 1995-August 2007.

shows why hourly wages decrease for hours above 50 in the left panel: earnings are close to flat above 50 hours.

Once again we are interested in how the relationship varies across subgroups. While our previous analysis allowed for age and education controls to shift wages, it did not allow them to interact with the shape of the wage-hours profile. To pursue this possibility we repeat the analysis when splitting the sample by age and education. Results are in the top panels of Figure 4. Note that the Figure does not provide any information about wage differences *across* age or education groups since each curve shows wages relative to the 40 hours bin for a given age or education, respectively. The main message from these two panels is that the same pattern noted for the overall male population also holds for each age and education group. Models of human capital accumulation and tournament models predict that young individuals might work long hours at low wages because of the future return to current hours. But significantly, Figure 4a shows that the cross-sectional patterns are effectively the same for young and old workers. The overall shape of the wage-hours profile is also the same across all education groups, though there is more variation across education groups than across age groups, see Figure 4b. The results for the wage-hours profile also hold when defining subgroups by both age and education, or splitting the sample by any of the covariates used in our baseline regression, as well as others such as number of children, spousal employment and hours, or recent unemployment experience.

The two middle panels of Figure 4 show the results when we repeat the same exercise across the census major occupation and sectoral categories. While there is heterogeneity across occupations and sectors, the key point is that for each of these occupations and sectors, the slope of the wage-hours profiles is positive below 40 hours, and negative after 45 or 50 hours. As noted earlier, all two-digit occupations and industries

Figure 4: Cross-Sectional Relationship between Wages and Hours by Different Groups, Across Datasets and Time



Sources and Notes: Figures (a) to (e) are based CPS ORG September 1995-August 2007. Figure (b) uses the same education classification as Figure 2b. Figure (c) and (d) use the 2010 major occupation (23) and industry (13) categories, and the dark wide line shows the line from Figure 3a. In Figure (e), the remaining datasets all report hours and wages for the previous year. The sample period for the CPS ASEC, PSID and NLSY is therefore 1996 through 2008. The ACS only starts in 2000 and is used through 2007 (in 2008 weeks worked last year are only available in bins). The SCF is conducted only every three years and features only a relatively small sample size, we use years 1992-2010 for the analysis here. For each dataset we use the set of control variables available in all datasets (a dummy for being black, being married, and a set of education and year dummies). Figure (f) reports result for the CPS ASEC because in the CPS ORG, information on whether earnings are imputed is only available from 1995 onwards. The years refer here to the reference year for wages and hours. Online Appendix Figure A.2 shows that from then onwards, the patterns in the CPS ORG and ASEC are very similar as also shown in (e).

also feature the heaviest concentration of individuals in the 40 hours bin.

The bottom panels in Figure 4 display results for two other exercises. Figure 4e repeats our basic analysis using several other data sets: the CPS ASEC (i.e. the March Supplement), the 2000 Census, the ACS, the PSID, the NLSY79 and the SCF. While details differ across these data sets in terms of the basic measures that we utilize, the figure shows that each of these generate not only the same qualitative shape for the wage-hours profile but also very similar quantitative properties.[4] Figure 4f shows that although the profile has changed between the 1970s and the present (see also Fuentes and Leamer, 2019), all of this change occurred prior to 1995 (see also Cha and Weeden, 2014); our estimated profile is very stable not only over the period 1995-2007 but also when we extend the analysis to the present. Importantly, the fact that earnings flatten beyond 50 hours is a stable feature of the data over the entire post 1970 period.

2.3.1 Wage vs. Salary Workers

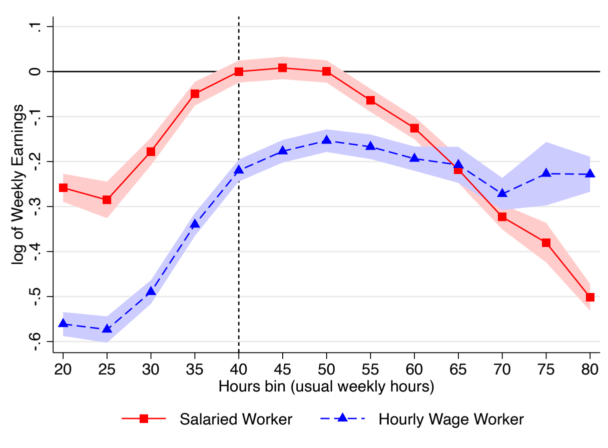
The previous analysis has shown that the hump-shaped wage-hours profile depicted in Figure 3 is robust to stratifying the data by a variety of observable characteristics. In this subsection we discuss one dimension along which the pattern is not robust: the distinction of wage versus salaried workers. Figure 5 shows the wage-hours profiles for each of these two groups. Differently than before, we normalize wages for both wage and salary workers relative to wages of salaried workers in the 40 hours bin. Figure 5 shows very different patterns for wage versus salary workers. In particular, whereas the curve for salaried workers displays a significant decline starting at 50 hours, the curve for wage workers is almost flat beyond 50, displaying only a very modest decline. The two curves also display somewhat different behavior in the range of 40 to 50 hours, with the curve for salaried workers being flat while the curve for wage workers exhibits a modest increase.

Part of these differences may reflect the impact of overtime pay regulations on compensation packages. Trejo (1991) presents evidence that jobs that are expected to have some periods with overtime hours paid at time and a half are associated with lower base wage rates. The logic is that the base wage rate adjusts to offset the effect of the overtime pay regulation so as to generate a particular level of expected compensation. A wage worker with current usual hours of 40 that expects to work overtime in the future will therefore have lower current wages than a similar worker that never expects to work overtime. This tends to depress observed wages for wage workers current usual hours of 40 relative to those with more than 40 hours.

But the more important issue for our purposes is that there are very few wage workers who actually work

[4]For the CPS ORG, the PSID and NLSY we have also estimated the relationship when removing individual fixed effects, i.e., we exploit within-person rather than cross-sectional variation. The same inverted U shape emerges, though the peak is now at 40 hours and the declines are somewhat steeper. These results are in Online Appendix Figure A.4.

Figure 5: Cross-Sectional Relationship between Wages and Hours by Job-Type



Source: CPS ORG September 1995-August 2007. *Notes:* The shaded areas are the 95% confidence intervals.

more than 50 hours, which is the region with the most striking difference. In particular, among workers who work long hours, 74% of them are salaried workers (compared to 47% in our overall sample). Put somewhat differently, for workers who are assessing their labor market opportunities conditional on working long hours, the vast majority of the opportunities they encounter would be salaried.

2.4 Measurement Issues

The previous subsection documented a striking fact: in the cross-section earnings are relatively flat beyond 50 hours, implying that wages fall significantly with usual hours worked beyond 50. One concern is that this decline in wages may be an artifact of measurement issues. In this subsection we summarize the results from several robustness checks and conclude that the decline in wages above 50 hours is not solely the result of measurement issues.

One possibility is that the comparatively flat earnings coefficients above 50 hours shown in Figure 3 are driven by top-coding. Top coding is of very minor importance for those with usual weekly hours below 45 hours, but does increase in importance with the level of usual weekly hours, increasing from just under 6 percent for those in the 50-54 hours bin to about 10 percent for those in the $60-64$ and $65-69$ hours bins. In Online Appendix B we show that datasets with different top-coding levels produce very similar wage-hours profiles in the long hours region, which suggests that top-coding is not of first-order importance.

A second possibility is that individuals in the long hours region are salaried workers who face temporary variation in hours but have a fixed salary. In this sense, the salary reflects compensation for expected hours rather than actual hours. If long hours individuals are disproportionately those with temporarily high hours, this would tend to flatten the earnings profile.[5] We assess this explanation using the small panel component of the CPS ORG to create a sample in which we have two observations per individual 12 months apart. Averaging across the two observations should dampen the effect of temporary variation in usual hours in the face of fixed compensation. However, when we run this alternative specification we find the same pattern quantitatively. We have also pursued this using the NLSY79 and PSID, which allows us to average over more years. Our main finding remains when we average over a five year period, see Online Appendix Figure A.3.

A third and related possibility is measurement error in hours. If people with high hours tend to be people who have over-reported their hours then this will show up as a negative effect of hours on wages. If classical measurement error was driving the results, then we would expect the averaging exercise just described above to produce very different patterns, which it does not. But this does not rule out measurement error stories that rely on non-classical measurement error; many long hours workers may consistently over-report hours. To assess this we use the linked observations between the CPS ORG and the American Time Use Survey (ATUS), which features information on usual weekly hours with a single observation on hours actually worked for a particular day. By pooling across individuals we can compute a synthetic measure of average weekly hours from the ATUS for individuals with reported usual hours in the CPS ORG within a particular hours bin.

Here we summarize the key findings; detailed results from this exercise are presented in Online Appendix C. First, in line with Frazis and Stewart (2014) we find that the two values track each other very closely for usual hours up to 70. We are not concerned about the larger deviation beyond 70 since there are very few individuals in that region. Second, while our measure of synthetic weekly hours computed from the ATUS is systematically lower than the reported measure in the CPS ORG and the gap grows as hours increase above 40, the difference remains relatively small, reaching around 5 hours per week in the 65-69 hours bin. In percentage terms, the gap is relatively constant in the 50-69 hours range, staying in the range of 6.5% to 8.5%.

While not insignificant, this discrepancy has little potential to change our key finding of flat earnings and decreasing wages above 50 hours. The reason is simple: if earnings are flat beyond 50 hours, then moving individuals from higher to lower hours bins does not affect earnings within a bin and hence will not affect average wages within a bin. Similarly, if hours are systematically overstated by a relatively constant percentage above 50, this will cause a uniform decrease in measured wages relative to true wages, but will not affect the slope of the wage-hours profile above 50 hours. In Online Appendix C we report results for

[5] Denning et al (2019) suggest that this explanation accounts for the low elasticity of wages to hours when using actual hours.

the case in which we adjust CPS hours based on the gap with reported hours in the ATUS. This adjustment has virtually no effect on the behavior of wages beyond 50 hours. In summary, taking reported hours from the ATUS time diaries at face value, our analysis of linked CPS ORG-ATUS data leads us to conclude that systematic over-reporting of usual weekly hours is not the dominant explanation for the relatively flat earnings beyond 50 hours.

Lastly, we have used the long panel feature of the NLSY79 to further cast doubt on the possibility that reported hours above 50 largely represent measurement error. As noted earlier, using the cross-section component of the NLSY79 we get essentially the same results that we found in the CPS ORG. We then use the panel component of the NLSY79 and find that individuals with high reported hours tend to have higher future wage growth, consistent with the evidence presented in Imai and Keane (2004) and Michelacci and Pijoan-Mas (2012).[6] If long hours were simply the result of individuals over-reporting their hours in a persistent fashion, then we would not expect to see that high hours are predictive of future wage growth.

In summary, while measurement error in hours certainly plays some role in shaping the observed profile of mean wages versus usual weekly hours, and will be included it in our later analysis, the evidence presented above strongly suggests that the relatively flat earnings in the long hours region and the resulting decline in wage rates is not purely a measurement artifact.

Our discussion to this point has focused on measurement issues concerning hours. There may also be concerns about measurement of labor income, in particular related to compensation that might not be included in reported earnings, e.g., deferred equity payments, employer contributions to health insurance or private pension plans. This information is unfortunately not available in many standard household data sets. The CPS ASEC does have information on employer contributions to health insurance and participation in employer-provided pension or retirement plans. Online Appendix Figure A.5 shows that employer-provided benefits to the degree we can measure them do not change our findings. More broadly, any lump-sum benefits would tend to make the wage-hours profile decrease even more steeply, while any benefits proportional to earnings, such as Social Security accruals, will have no effect at all.[7]

2.5 Determinants of Hours Worked

In the next section we will present a heterogeneous agent model that we use to understand the cross-sectional patterns in wages and hours. To guide the choice of which dimensions of heterogeneity to include in our

[6] See Gicheva (2013) and Barlevy and Neal (2019) for evidence on dynamic effects in the context of professional labor markets.

[7] Eisfeldt et al (2021) document the increased importance of deferred equity payments using Compustat data, though they cannot link payments to individual workers. We do not think this exerts a large effect on the slope of our wage-hours profile in the long hours region. The reason is that as previously documented, the profile is stable over time whereas Eisfeldt et al (2021) document a large increase in this form of compensation in the 1990s.

model specification, this subsection assesses the empirical significance of various factors in accounting for the dispersion in hours and the incidence of long hours. We report three main findings. First, controlling for age and education, other demographic variables and non-labor income (reflecting all household income other than the individual's own labor earnings including capital income) have very little explanatory power in accounting for either the dispersion in hours worked, or the incidence of long hours. Second, adding lagged hours to the specification increases the explanatory power by more than an order of magnitude. Interpreting lagged hours as a proxy for persistent unobservable differences in preferences, this suggests that preference heterogeneity is a key driver of heterogeneity in usual hours of work. Third, while three digit occupation and/or industry controls do have modest explanatory power when only including contemporaneous variables, their explanatory power drops significantly when lagged hours are included.

We begin by examining the relationship between usual weekly hours, demographic characteristics and non-labor income in the CPS-ASEC since it provides us with a more extensive measure of income from other sources than the CPS-ORG. For these exercises we stratify the population of males aged 25-64 by education (less than high school, high school and some college, bachelor's degree college, more than a bachelor's degree). For observables we include age (four groups, 25-34, 35-44, 45-54, 55-64), race/ethnicity (three groups: black, hispanic, and non-black, non-hispanic), marital status (married, not married), number of own children aged 0-4 (0, 1, 2, 3+) and aged 5 and older (0, 1, 2, 3+) in the household, and quintile for non-labor income (i.e., all household income less the individual's own earnings). Data is pooled over the survey years 1996-2008, so that earnings and weekly usual hours correspond to the years 1995-2007. Our main exercise runs OLS regressions of usual weekly hours on these explanatory variables. We also run a linear probability specification in which the dependent variable is whether the individual has usual weekly hours of at least 50. Because this exercise yielded the same message as the OLS exercise, we focus here on the results of the OLS exercise; results of the linear probability specification are contained in Online Appendix A.2.1.

Our primary interest is the explanatory power of the observables. For this reason, estimated coefficients from the OLS exercise are contained in Tables A.1 in the Online Appendix, and here we focus on the results in the first four columns of Table 1, which indicate the R^2 for the full regression with all of the dependent variables included, as well as the R^2 when we eliminate various individual observables. The key message is that observables explain a very small fraction of the cross-sectional variation in hours worked. A secondary message is that the very modest overall explanatory power of these observables does not seem to be driven by any particular observable.

Next, we repeat the previous analysis, but now using the small panel component in the CPS-ASEC, which provides observations for two subsequent years per individual. To the extent that persistent hetero-

Table 1: Explantory Power of Different Determinants of Log Usual Weekly Hours Worked in the CPS ASEC

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.019	0.027	0.028	0.023	0.289	0.304	0.336	0.407
Excluded Regressors								
Age	0.014	0.021	0.021	0.016	0.285	0.302	0.333	0.402
Race/Ethnicity	0.016	0.023	0.026	0.022	0.287	0.303	0.336	0.407
Marital Status	0.014	0.020	0.023	0.019	0.283	0.303	0.335	0.406
Age & # of Children	0.018	0.026	0.026	0.022	0.288	0.303	0.336	0.406
Other Income Quintile	0.016	0.024	0.025	0.021	0.286	0.303	0.335	0.406
Lagged Hours	–	–	–	–	0.030	0.024	0.030	0.027

Source: CPS ASEC 1996-2008. *Notes:* The first row of the reports the R^2 of an OLS regression with log usual hours worked as the dependent variable and including all regressors stated in the main text. Rows 2-6 reports the R^2 of the regression the set of regressors reported in the first column is dropped. The right panel uses the subsample of individuals, who took the CPS ASEC in two subsequent years and uses the second observation for each individual. Lagged hours as control variable are taken from the first observation for each individual. The education classification is the same as in Figure 2b.

geneity in preferences is a key driver of heterogeneity in usual hours of work, lagged hours of work would be a proxy for persistent differences in preferences, and so we include this in the regressions. The coefficient estimates are again included in the Online Appendix Table A.1, and the results for the R^2 are reported in the last four columns of Table 1. Comparing across the left and right panels of Table 1 yields a simple message: the R^2 increases by more than an order of magnitude when we include lagged hours.[8] These results are consistent with an important role for preferences as a non-wage driver of hours variation.

We have also repeated this analysis including three-digit occupation controls. The key findings are as follows. If we add controls for occupation without including lagged usual weekly hours, then the R^2 increases by about 0.10 on average across specifications, see Online Appendix Table A.3a. This is about one-third as large as the increase when we add lagged hours. However, conditional on including lagged usual weekly hours, the increase in the R^2 from adding controls for occupation is less than 0.05 for three of the four education groups. We conclude that while occupation contains predictive power for hours, it is a relatively unimportant driver of the dispersion in hours once one controls for past hours. We have also repeated this exercise using three digit industry controls, and the results are similar.

One limitation of the CPS-ASEC data is that it does not include information on assets. In order to assess the role of this variable as an explanatory variable we turn to data from the SCF, which is thought to have the highest quality data on wealth. We redo the pure cross-section exercise as above (lagged hours are not

[8]The left panel in Table 1 has a larger sample size than the right panel. The same message emerges if we redo the left panel using the smaller sample, see Table A.4. In Appendix A.2.2 we report the R^2 results for the smaller sample for all specifications based on the ASEC discussed in this section.

Table 2: Explantory Power of Different Determinants of Log Usual Weekly Hours Worked in the SCF

	LHS	HS	Bach	Bach+
All Regressors	0.041	0.032	0.034	0.024
Excluded Regressors				
Age	0.038	0.030	0.030	0.024
Race/Ethnicity	0.035	0.022	0.028	0.024
Marital Status	0.041	0.031	0.029	0.020
Age & # of Children	0.037	0.029	0.032	0.022
Other Income Quintile	0.032	0.022	0.027	0.019
log Wealth-Income Ratio	0.038	0.031	0.034	0.023

Source: SCF 1992-2010. *Notes:* The first row of the reports the R^2 of an OLS regression with log usual hours worked as the dependent variable and including all regressors stated in the main text. Rows 2-6 reports the R^2 of the regression the set of regressors reported in the first column is dropped. The education classification is the same as in Figure 2b.

available in the SCF), but now also include wealth relative to labor earnings as an independent variable. This allows us to control for the fact that holding all else constant, individuals with higher earnings would be expected to accumulate higher wealth both for life-cycle saving (i.e., retirement) and for precautionary saving motives. As before, here we focus on the R^2 for the OLS specification, with the coefficient estimates and results for the linear probability specification reported in the Online Appendix in Tables A.8 and A.9 respectively. The key results are presented in Table 2 and show that the R^2 is low, typically around 0.03, broadly consistent with the cross-sectional results using the CPS. The wealth to earnings ratio does provide a small amount of explanatory power, but is comparable to the explanatory power of the other observables in the CPS analysis.[9]

In summary, the main take-away is that contemporaneous observables have very little power in explaining hours worked or the probability to work long hours. The most important determinant is lagged hours worked, which plausibly proxies for persistent differences in tastes for work. Structural models of male labor supply have long included preference heterogeneity to account for dispersion in hours worked for workers with similar observable characteristics. (See, for example the early analyses of MaCurdy (1981) and Altonji (1986).) The reduced form empirical analysis of Abowd and Card (1989) established the importance of non-wage factors as a dominant source of movements in hours in panel data. Recent work that examines the cross-sectional properties of labor supply (e.g., Heathcote et al (2014), Erosa et al (2016), and Boerma and Karabarbounis (2021)) attributes an important role to heterogeneity in preferences, and the findings presented above are consistent with this.

[9]We found similar results when running this regression separately by age group (see Appendix A.2.4), and using log wealth rather than the log wealth to income ratio (see Appendix A.2.5).

3 A Structural Model of Hours and Wages in the Cross-Section

The previous section documented two salient and robust features of the joint wage-hours distribution in the US: (i) usual weekly hours are heavily concentrated at 40, with a long upper tail, and (ii) mean wages are hump-shaped in hours of work, increasing until 50 hours per week and decreasing thereafter. In this section we develop and estimate a static model of labor supply along the intensive margin that can account for these facts both qualitatively and quantitatively. Our model nests a standard textbook model of labor supply and we show that this standard model is unable to account for these features of the data.

Our choice to work with a static model is motivated by two considerations. First, it facilitates a more transparent examination of the key forces at work. Second, as documented in Section 2, the key facts that we study are very stable across age groups, suggesting that dynamic life cycle forces such as human capital accumulation are not playing a dominant role. This modeling choice will in turn affect our estimation strategy. In particular, our benchmark estimation exercise will focus on males aged 50-54. We focus on males because our analysis abstracts from the participation margin, and this margin is arguably less important for males. Our choice of an age group balances the desire to have a sample for which extensive margin considerations due to early retirement are not too important at the same time that the potential dynamic returns to working additional hours are less relevant. The human capital literature typically assumes that individuals older than 50 face very low returns to human capital accumulation, see for example, Heckman et al (1998).

3.1 Model

Our benchmark model assumes individuals are heterogeneous along two dimensions: productivity and preferences. The analysis in the previous section highlighted the limited power of observable characteristics to explain hours of work, and consistent with this we will treat preference heterogeneity as a single unobservable factor that influences hours. In a later section we extend our analysis to include heterogeneity in non-labor income (reflecting, for example, spousal labor earnings and income from capital) and show that this source of heterogeneity is relatively unimportant for our findings. This is consistent with the results from the previous section showing that other income has relatively little explanatory power for hours in the cross-section data, and motivates our choice to abstract from it in our benchmark model.

There is a unit mass of individuals with preferences over consumption and hours of work given by:

$$\frac{1}{1-(1/\sigma)}c_i^{1-\frac{1}{\sigma}} - \frac{\alpha_i}{1+(1/\gamma)}h_i^{1+\frac{1}{\gamma}}$$

Individuals are heterogeneous in terms of preferences, captured by the parameter α_i, and productivity, which is denoted by z_i. While we assume a single dimension of preference heterogeneity, we emphasize that heterogeneity in α should be interpreted as summarizing the net effect of the many factors that might influence the relative marginal utility of consumption and time devoted to market work. We assume that α and z are jointly log normally distributed, characterized by the mean and standard deviation of $\log\alpha$, the mean and standard deviation of $\log z$, and the correlation between $\log\alpha$ and $\log z$, which we denote by μ_α, σ_α, μ_z, σ_z, and $\rho_{\alpha,z}$ respectively.[10] The two preference parameters σ and γ are the same for all individuals.

The novel feature of our model is that individuals face a nonlinear schedule for earnings as a function of hours. To capture this we write the budget equation for individual i as:[11]

$$c_i = z_i E(h_i)$$

In what follows we will refer to $E(h)$ as the *earnings function*. We also define the *wage function*, $W(h)$ as:

$$W(h) = \frac{E(h)}{h}.$$

We want to emphasize that due to potential selection issues, these theoretical objects are distinct from the cross-sectional relationships between hours and earnings or hours and wages that one observes in the data. In what follows we will use the terms *wage function* and *earnings function* to refer to these theoretical objects, and the terms *equilibrium wage profile* and *equilibrium earnings profile* to refer to the empirical cross-sectional relationships induced by worker choices.

The standard textbook model with a linear budget equation corresponds to the case in which $E(h) = h$, where we suppress the wage per efficiency unit of labor, w, since in our partial equilibrium exercise it can be subsumed into the mean of z_i without loss of generality. This specification implies a unitary elasticity of earnings with respect to hours. A second case of interest, due to French (2005), is $E(h) = h^{\bar{\theta}}$. This imposes a constant elasticity of earnings with respect to hours but does not restrict it to equal unity. We generalize the model of French (2005) by allowing this elasticity to vary across regions of the hours distribution. In particular, whereas French (2005) assumed a log linear relationship between earning and hours, we assume a piecewise log-linear relationship:

$$E(h_i) = A(h_i) h_i^{\theta(h_i)}$$

where $\theta(h)$ is a step function and the $A(h)$ term is included purely to maintain continuity of the earnings

[10] We also experimented with a Pareto log normal distribution as in Badel et al (2019), which has more mass in the right tail, but found that this had little effect on our key results.

[11] Our benchmark model abstracts from taxes. Introducing progressive income taxation as in Heathcote et al (2014) leaves our results effectively unchanged.

function at a point of discontinuity in the $\theta(h)$ function.[12] An appealing feature of our functional form for earnings is that the function $\theta(h)$ offers a clear and flexible mapping for how the elasticity of earnings with respect to hours varies across the hours distribution.

It is intuitive that this generalization of French (2005) might help to account for the two key properties of the cross-sectional wage and hours data that we highlighted earlier. First, a kink in the earnings function associated with a downward jump in $\theta(h)$ will tend to generate bunching in the hours worked distribution at the kink. Second, kinks in the earnings function will necessarily impact the shape of the observed wage-hours profile. In what follows our goal is to assess the extent to which this extension helps us to account for the patterns in the data, and if so, what it implies for the properties of the $\theta(h)$ function.

3.2 Measurement Error

We argued earlier that the non-monotonic wage-hours profile is not purely a reflection of measurement error. However, we do want to allow for the possibility that measurement error plays some role and so will assume that hours are measured with error.

In our benchmark exercise we allow for measurement error in log hours that is classical subject to one qualification. The qualification is that if an individual has true hours equal to 40 we assume that they do not report with measurement error. The rationale for this is intuitive – it is virtually impossible to generate a large spike at 40 hours if we assume that everyone reports hours with classical measurement error. In fact, there is good reason to believe that measurement error more likely serves to increase the spike at 40 rather than diminish it, since another feature of the reported usual hours distribution is that there is heaping at all values ending in either a zero or a five. A natural interpretation is that individuals tend to round to a multiple of five when reporting usual weekly hours. We do not attempt to incorporate this type of measurement error, which partly motivates our decision to focus on five hour bins when we connect our model to the data. For those who do not work exactly 40 hours we assume that log hours are reported with normally distributed measurement error that is iid across individuals with mean zero and standard deviation σ_m.

In contrast to measurement error in hours, classical measurement error in log earnings has relatively little impact on our findings. Within an hours bin, this type of measurement error has no impact on the average log earnings in the bin and little impact on average log wages. Classical measurement error in earnings does impact the overall correlation between wages and hours, but for reasonable values of measurement error this

[12] That is, the function $A(h)$ is constant over any interval in which $\theta(h)$ is continuous; as a normalization we impose $A(0) = 1$. It seems natural to impose continuity given that our empirical analysis did not suggest any significant discontinuities in the empirical earnings-hours profile.

effect is small. For this reason we abstract from measurement error in earnings.[13]

3.3 Moment Matching Exercise

In this subsection we describe the procedure we use to estimate the parameters of our model. The basic strategy is to choose values for our model parameters so that the model best matches a key set of empirical moments that characterize the cross-sectional distributions of weekly hours and wages.

3.3.1 Data

We estimate our models for males aged 50-54 for the reasons discussed earlier: this is a group with strong labor force attachment for whom dynamic considerations are likely to be relatively unimportant. Since we will also stratify our sample by education, we will use moments from the ACS data set. Relative to the CPS-ORG it offers a larger sample size, a higher threshold for top-coding, and top-coded values that represent the average earnings of those subject to top-coding rather than the top-code itself.[14]

3.3.2 The Earnings Function

We parameterize the earnings elasticity step function $\theta(h)$ to have three regions: $\theta(h) = \theta_s$ for short hours (i.e., h below 40), $\theta(h) = \theta_m$ for hours in the interval $[40,50)$, and $\theta(h) = \theta_l$ for long hours (i.e., h greater than or equal to 50). The choice of 40 hours for a kink in the log earnings function is empirically motivated and a priori intuitive since workers will tend to concentrate their hours at a kink in the earnings function. Allowing for a second kink at 50 hours is also perhaps intuitive given that the empirical wage profile changes slope at this point. Additionally, we found that including this second kink was important to generate sufficient concentration of workers with 50 or more hours in the 50-54 hours bin. The data suggests that one might want to include an additional region for hours below 30. It would be relatively simple to do this, but given that our current application is based on data for males and there are so few males below 30 hours, we have chosen to not focus on that region and thereby reduce the set of parameters. Moreover, as explained later, our primary focus is on the properties of hours and wages in the long hours region.

An appealing feature of our earnings function specification is that it allows for different hours-earnings trade-offs for workers desiring part-time work schedules than for workers desiring longer work schedules.

[13] Heathcote et al (2014) estimated no measurement error in earnings, which they argued was consistent with other results in the literature.

[14] In Section 2 we documented that in the aggregate the above moments are very similar across several datasets. In Section 2.4 and Online Appendix B, we document that this strategy of replacing top-coded earnings with the average earnings of top-coded individuals largely eliminates the role of top-coding in our context.

While our choice of $\theta(h)$ imposes quite a bit of structure, we will see that it is sufficiently flexible to account quite well for the features of the data that we target.

Our specification for $\theta(h)$ and preferences implies that all individuals will work positive hours, so there is no selection of individuals into employment. Introducing fixed costs as in Cogan (1981) or altering the shape of the earnings function at low hours as in Prescott et al (2009) would allow us to generate an active extensive margin, but given our application to workers with very high participation rates, this is not a first order issue and we not pursue it in this paper.

To illustrate the role of our specification for the earnings function in allowing us to match the data, we also estimate two restricted versions of our model. The first of these is the textbook linear earnings model in which $\theta_s = \theta_m = \theta_l = 1$. The second is the French model with log linearity, i.e., $\theta_s = \theta_m = \theta_l = \bar{\theta}$ and $\bar{\theta}$ can differ from unity. In what follows we will refer to the linear model as Model 1 (M1), the log linear model of French (2005) as Model 2 (M2) and our piecewise log linear model as Model 3 (M3).[15]

Our approach focuses on understanding the patterns in the data from a pure labor supply perspective; we assume that each individual freely chooses their hours of work, taking as given the opportunities reflected by $E(h)$. That is, if we observe an individual who works 40 hours, we assume that the wages being offered for other levels of hours were such that the individual preferred to work 40 hours, not that the individual did not have the option to work a different number of hours. To the extent that firms do not desire to hire workers for a particular level of hours, this will manifest itself as low wages associated with that level of hours. That is, our earnings function embeds factors that operate on the firm side and affect the demand for different workweeks. Our earnings function should be interpreted as the opportunities that the worker faces in the market more broadly and not necessarily the options available at a given firm. Altonji and Paxson (1988) emphasized workers seeking to change their usual weekly hours as a source of turnover at the firm level.

3.3.3 The Role of Our Distributional Assumptions

It is important to highlight the role of our distributional assumptions. If we placed no restrictions on the joint distribution of α_i and z_i then the linear model (M1) can perfectly match any cross-section data on hours and wages since there are two data values and two degrees of freedom for each individual. Given the high concentration of workers with $h = 40$ in the data, this procedure would generate spikes in the distribution of tastes for work, and the position of the spike would need to be conditioned on any other variables that

[15] We have also considered a model in which θ_s is unrestricted but $\theta_m = \theta_l = 1$. The key message is that for empirically reasonable values of θ_s, this model offers a slight improvement relative to the linear model but performs markedly worse than our benchmark model.

influence hours of work. We find this specification to be unappealing and so instead restrict attention to heterogeneity that is lognormally distributed. While this choice will clearly influence our inference, we think it is of interest to ask whether one can match the cross-section evidence subject to this constraint. We note that our distributional assumption is common in the literature, and in particular is the one adopted by Heathcote et al (2014).

The distribution of worker characteristics α and z that we estimate using data for older males should be interpreted as reflecting any history dependent evolutions. In particular, our specification is fully consistent with the possibility that choices about hours of work when young had effects on future productivity, and that our taste shifter evolves over time. Our key assumption is that forward looking effects of labor supply on wages are of second order importance for individuals aged 50-54.

3.3.4 Externally Calibrated Parameters

Our estimation exercise fixes the values of σ and γ. Our benchmark results are for the case in which σ tends to one, implying offsetting income and substitution effects, and $\gamma = 0.50$. Neither of these values are substantively consequential for our results. Changing the value of σ induces a correlation between z and h holding α constant. In our moment matching exercise this effect is essentially undone by changing the value of $\rho_{\alpha,z}$, leaving the overall model fit and the estimates of the other parameters, specifically the θ_j, unchanged. The value of γ is not important for our estimation exercise because changes in γ will essentially be undone by changes in the standard deviation of the unobserved preference heterogeneity.

3.3.5 Estimated Parameters

Our choices to this point leave nine parameter values not yet assigned: the five parameters for the joint distribution of α and z (μ_α, σ_α, μ_z, σ_z, and $\rho_{\alpha,z}$), the three θ_j values that define the earnings function, and σ_m, the standard deviation of classical measurement error in log hours. The choice of μ_z is effectively a choice of units, and so we normalize it to zero. We choose values for the remaining eight parameters in order to fit the following 18 moments: the standard deviation of log wages, the distribution of hours worked across five hour bins between 30 and 69 hours as well as the fraction with hours greater than or equal to 70 (9 moments), and the equilibrium wage profile in logs across five hour bins between 30 and 70 hours (8 moments).[16] Our measure of fit is the sum of squared errors.

[16] For our estimating sample of males aged 50-54 only 3 percent of the observations lie outside of the $30-70$ hours range, which is why we do not seek to include wages for those workers in the moment matching exercise. As noted earlier, if we wanted to match wages for those with hours below 30 we would need to include a fourth region for the step function $\theta(h)$.

In implementing this procedure, we found that the loss function was fairly flat for a range of values for θ_s, the parameter capturing the marginal return to additional hours in the part-time region. This range covers the interval $[1.4, 1.9]$. We resolve this issue by drawing on the relatively extensive literature on the part-time wage penalty. In particular, Aaronson and French (2004) argued that Social Security rules create a setting in which one can identify the effects of exogenous reductions in hours for older males. Their estimate implies $\theta_s = 1.4$, which lies at the lower end of the flat region of our loss function. Because our primary focus is on outcomes above 40 hours, we choose to impose the value from Aaronson and French as our baseline. Our key results are robust to choosing θ_s in the interval $[1.4, 1.9]$.

4 Results

This section reports our results in four steps. In the first step, we estimate the parameters of our model on the subsample of males aged 50-54 with a high school degree. We choose males aged 50-54 for the reasons discussed earlier: this is a group with strong labor force attachment for whom dynamic considerations are likely to be relatively unimportant. We stratify by education to allow for the possibility that the earnings function will vary across groups, and choose the high school group as our benchmark sample because it is by far the largest. The important finding from the first step is that our estimated model provides a good fit to the data.

In the second step we contrast the ability of our model to fit the data relative to the restricted models M1 and M2 described earlier. We show that these two restricted models fail to capture salient and robust features of the cross-sectional hours and wage distributions that we highlighted in Section 2.

In the third step, we extend our analysis to two other educational subgroups of males aged 50-54: those with a bachelor's degree and those with more than a bachelor's degree. (We do not study the less than high school group, since this group has a much lower participation rate and our model abstracts from participation.) To highlight the role of the earnings function, we ask to what extent our model can explain the different patterns across educational subgroups by varying only the properties of the earnings function and holding all other parameters fixed. We find that this exercise generates a good fit to the data for these two educational groups.

In the fourth step, we ask whether our model estimated on data for workers aged 50-54, for whom dynamic considerations such as human capital accumulation are likely to be less important, can also account for the patterns in the data for younger individuals. To the extent that dynamic considerations may be more important for younger workers, there is no presumption that our model estimated on older individuals will fit

Table 3: Estimated Parameters, Males 50-54 With High School Degree

μ_α	σ_α	σ_z	$\rho_{\alpha,z}$	σ_m	θ_s	θ_m	θ_l
-13.64	1.61	0.589	-0.33	0.04	1.40	0.058	0.034

the patterns equally well for younger workers. Nevertheless, we find that our estimated earnings functions also generate a good fit to the data for this younger group.

4.1 Benchmark Estimates

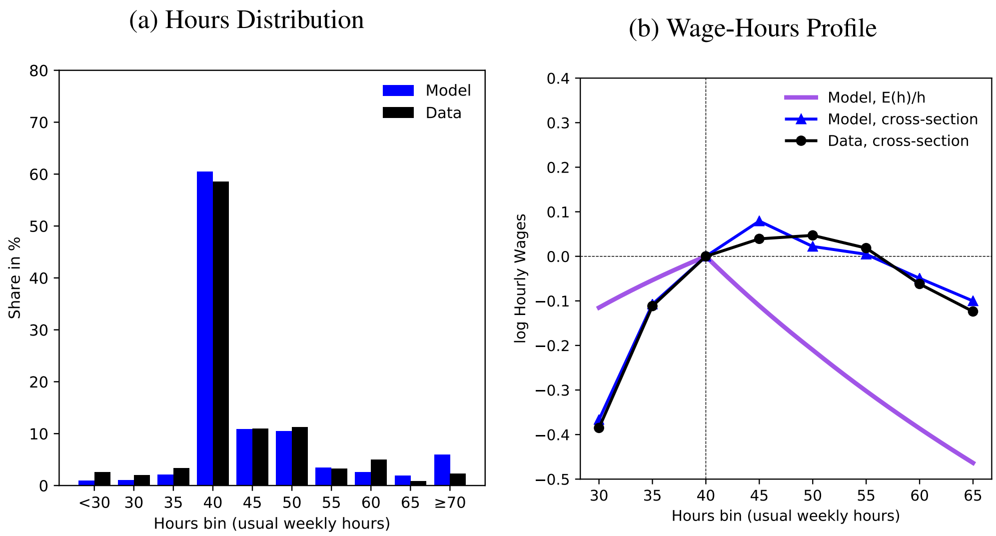
We first present and discuss the estimated parameter values, and later consider the fit of our estimated model. Table 3 presents benchmark parameter estimates when applying our moment matching exercise to the sample of males aged 50-54 with a high school education. Our primary focus is on the estimates for the θ_j. Since θ_s has been set to 1.4 (the value estimated by Aaronson and French (2004)), we focus on the values of θ_m and θ_l. Both of these values are smaller than one, indicating a "long hours" penalty, with the penalty being marginally higher in the region above 50 hours.[17] The estimated values of θ_m and θ_l are not that different from one another, and the deterioration in fit is not that large if they were restricted to be equal; the main effect is that the model underpredicts the mass in the 50-54 hours bin in that case. While the kink at 50 is not particularly important for the high school group it is more important for the bachelor and bachelor plus groups that we study later. The size of the penalty for working more than 40 hours is large: a worker who increases hours from 40 to 50 will experience an earnings increase of only about 1.3 percent, implying a decrease in hourly wages of almost 20 percent.

Selection effects in our estimated model are driven by the correlation between α_i and z_i. The estimated value of $\rho_{\alpha,z}$ is negative, indicating that high productivity individuals tend to be more willing to work longer hours, thus implying that high hours workers will be positively selected with regard to productivity. The sign of this correlation determines where the model implied equilibrium earnings profile lies relative to the earnings function $E(h)$. In particular, if $\rho_{\alpha,z}$ is less than (greater than) zero then the equilibrium earnings profile will lie above (below) the earnings function in the $h > 40$ region; the relationship between these two profiles will be the reverse in the region below 40 hours. In the special case of $\rho_{\alpha,z} = 0$ the two are identical modulo any measurement error.

We offer a heuristic explanation for why our estimation generates a negative value for this correlation. Generating the large concentration of individuals in the 40-44 hours bin requires a large kink in the earnings

[17] Importantly, this property holds for any value of θ_s within the interval $[1.4, 1.9]$ over which our loss function was relatively flat. Increasing θ_s increases the estimates of both θ_m and θ_l, but even if we set $\theta_s = 1.9$ the implied values for θ_m and θ_l are 0.211 and 0.123 respectively, and so are still far below unity.

Figure 6: Fit of 3-Region Model (M3): High School



function at 40. Holding the size of the kink constant, one can pivot the earnings function around the $h = 40$ point to vary the gap between the the equilibrium earnings profile in the data and the earnings function in the long hours region. As this gap changes, the required value of $\rho_{\alpha,z}$ also changes. In particular, as we pivot $E(h)$ counter-clockwise around $h = 40$, the earnings function $E(h)$ rotates upward in the region above 40 hours, changing its position relative to the equilibrium earnings profile in the data and hence increasing the implied value for $\rho_{\alpha,z}$. Because this rotation holds the size of the kink at 40 constant, it implies higher values for each of θ_s, θ_m, and θ_l. We make two remarks. First, the rotation that would require a positive value for $\rho_{\alpha,z}$ requires an implausibly large value for θ_s. Second, this rotation is not neutral with respect to the model's ability to fit other moments, and it turns out that the overall fit of the model is much worse when one imposes a positive value for $\rho_{\alpha,z}$.[18]

A negative value of $\rho_{\alpha,z}$ is intuitive. Models of human capital investment suggest a strong theoretical rationale for a negative correlation between productivity and tastes for leisure. First, since the return to human capital investment is increasing in future labor supply, individuals who expect to work more have greater incentive to invest, holding all else constant. Second, if part of the opportunity cost of human capital investment is foregone leisure, a lower value for leisure would also lead one to do more investment in human capital. Other factors could also generate this correlation; for example, a negative health shock might simultaneously lower productivity and increase the disutility associated with work.

We now turn to the issue of model fit. Figure 6 shows that our model fits the data for both the hours

[18]We noted earlier that our loss function was relatively flat for θ_s in the interval $[1.4, 1.9]$. Consistent with the heuristics just described, this is associated with a range of values for $\rho_{\alpha,z}$, but importantly this range is $[-.32, -.25]$, which is both relatively small and entirely in the region $\rho_{\alpha,z} < 0$.

distribution and the equilibrium wage profile. In particular, it captures three key features: (i) the hours distribution is heavily concentrated in the 40-44 hours bin, (ii) there is a significant mass of workers that work 50 or more hours, and (iii) the equilibrium wage profile is non-monotonic. Online Appendix Figure D.1 shows that the model also closely replicates the standard deviation of log wages across the hours distribution, and well as mean hours and the standard deviation of hours across the wage distribution.

Panel (b) also includes the estimated wage function. Consistent with our previous heuristic discussion, the wage function lies below the equilibrium wage profile in the region above 40 hours, and above it in the region below 40 hours. The large kink in the wage function at 40 hours is readily apparent; while the kink at 50 hours is much smaller and thus not visible given the current scale, we do want to remind the reader that there is a smaller kink at 50 that does affect the mass of workers in the 50-54 hours bin.

By way of summary, we highlight two key results from our estimation. First, the earnings technology is highly non-linear: the elasticity of earnings with respect to hours is above one for those working less than 40 hours per week, but is far below one for those working 40 hours per week or more. In later sections we show that this result has several important implications. Second, selection on unobservables plays an important role in shaping the cross-sectional relationship between wages and hours. In particular, one cannot infer the choices that an individual worker faces simply by looking at the cross-sectional relationship between wages and hours.

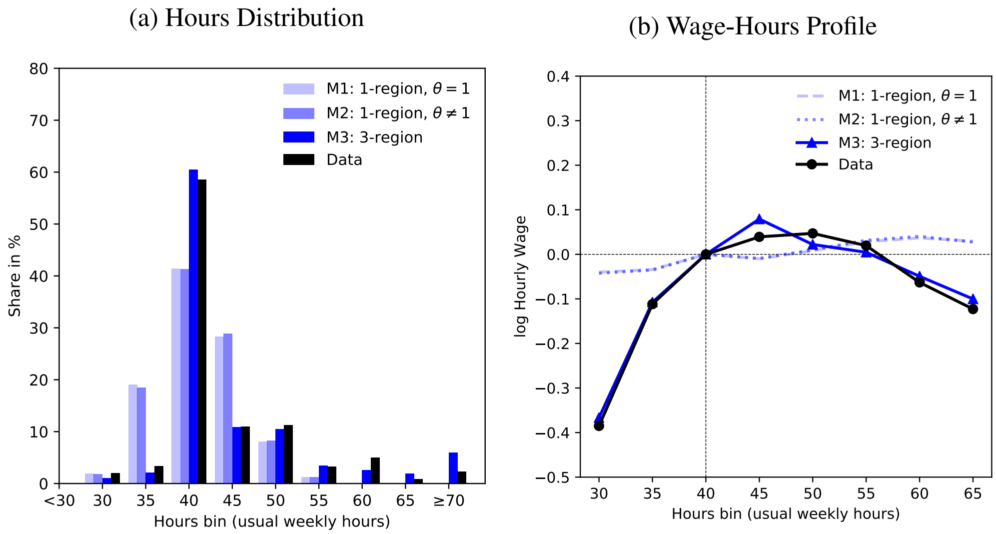
4.2 Comparison With Other Models

In this subsection we compare the ability of our model to account for the key patterns in the data relative to the two simpler models, M1 and M2. Recall that M1 is the textbook linear earnings model with all three θ_j set to unity, and M2 is the French (2005) specification in which the θ_j are all equal but not necessarily equal to unity.

Panel (a) of Figure 7 compares the fit to the hours distribution, and Panel (b) compares the fit to the equilibrium wage profile. Models M1 and M2 produce virtually identical results; allowing for an elasticity larger (smaller) than one introduces a positive (negative) correlation between hours and wages holding everything else equal, but in the estimation this is effectively undone by changing the extent of correlation between α and z. Perhaps not surprisingly, the loss function for M2 is fairly flat over $\bar{\theta}$ values.

Most importantly, both M1 and M2 do not capture the salient features noted earlier: both fail to produce sufficient concentration in the 40-44 hours bin, while at the same time failing to generate sufficient mass above 50 hours, and neither can generate the non-monotonic pattern in the data. We note that if one were to focus purely on standard first and second moments of the wage and hours distribution, then Model 1 (and

Figure 7: Comparing Fit of Models M1-M3: High School



hence also Model 2) is able to perfectly replicate the data; its shortcomings are only made apparent when one considers the richer set of moments that are the focus of our analysis. The key message is that our relatively simple generalization yields a very substantial improvement in the ability of the model to capture the salient features of the data.

4.3 Males 50-54 With More than High School

In this subsection we ask whether our model can also account for the cross-sectional patterns for the two other subgroups of males aged 50-54: those with bachelor degrees and those with more than a bachelor's degree.

One strategy would be to re-estimate all of the model parameters for these two educational attainment groups. This would implicitly allow the nature of unobserved heterogeneity as captured by μ_α, σ_α, and $\rho_{z,\alpha}$ to account for the different patterns across educational groups. We instead carry out an exercise that highlights the novel feature of our analysis, which is the connection between the non-linear earnings function $E(h)$ and the hours distribution. Specifically, we hold all parameters from the baseline estimation fixed except for θ_m and θ_l. For each group we set these two values to match two moments from the hours distribution: the fraction of individuals in the 40-44 hours bin and the fraction of individuals in the 50-54 hours bin. We then assess the ability of this specification to account for the broader features of the data, specifically the equilibrium wage profiles, which are not targeted.

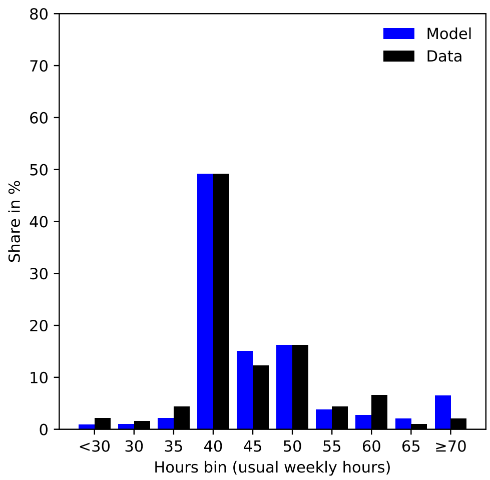
Table 4 shows the implied values for the θ_j for both our baseline model and the two other education groups. The implied values of the θ_j are as expected given the patterns that we noted in Section 2; because

Table 4: θ_j Values for Males Aged 50-54

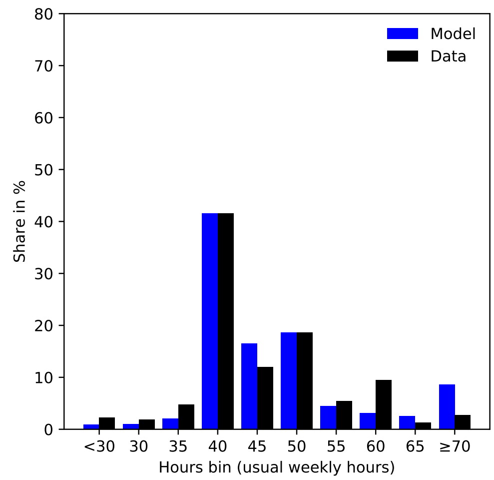
	θ_s	θ_m	θ_l
High School	1.40	0.058	0.034
Bachelor	1.40	0.092	0.037
Bachelor +	1.40	0.126	0.047

Figure 8: Fit of 3-Region Model (M3): More than High School

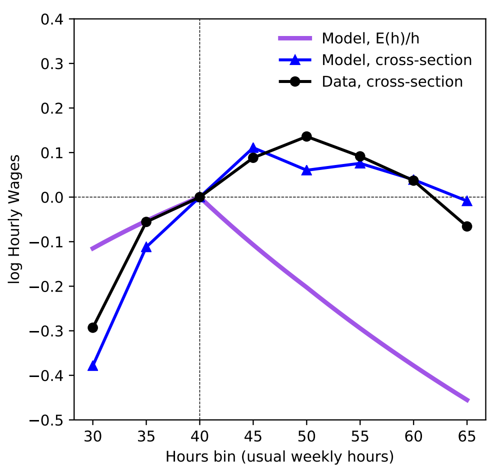
(a) Bachelor: Hours



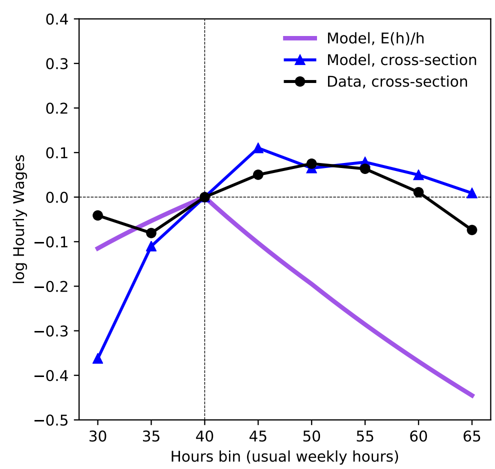
(b) Bachelor+: Hours



(c) Bachelor: Wages



(d) Bachelor+: Wages



the fraction of individuals in the 40-44 hours bin is decreasing in educational attainment, the required kink at 40 is also decreasing in education, implying that θ_m is increasing in education. Holding θ_l fixed, these changes in θ_m create larger kinks at 50 hours, and would create excessive concentration in the 50-54 hours bin. For this reason the values of θ_l also increase with education.

While our estimated values of the θ_j are pinned down by the hours distribution, the implication that both θ_m and θ_l increase with education would seem to conform with conventional wisdom about the return to long hours and educational attainment. Importantly though, the long hours penalty remains substantial even for the bachelor plus group, with a wage decline of almost 18 percent for an individual that increases hours from 40 to 50.[19]

Figure 8 shows the fit of the model for these two groups. Not too surprisingly given our estimation strategy, the model accounts well for the hours distributions for both groups. In both cases the model implies that there is a slight excess of workers in the 70+ bin; while we do not pursue it further here, we could introduce additional curvature in the log earnings function by adding another kink at 60 to address this. The model also does well in capturing the non-monotonicity in the equilibrium wage profiles. The fit is not as good as it was for the high school group, but recall that we are holding all parameters fixed except for θ_m and θ_l. Nonetheless, the estimated profiles do track key features of the data.[20] The bottom panels of Figure 8 also display the wage functions, and as before we note that while the large kink at 40 hours is readily apparent, the smaller kink at 50 is not visible at this scale, though it does exert an effect on the hours distribution.

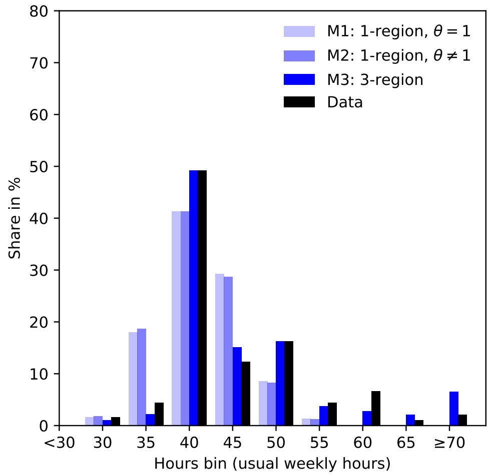
To contrast the fit of our model with the two simpler specifications, we have estimated both Models M1 and M2 for each of these two education groups. Importantly, for this exercise we estimate all the model parameters for M1 and M2, whereas for M3 only θ_m and θ_l are re-estimated. Figure 9 compares the performance of these models with the specifications in Figure 8. The basic message mimics our earlier results for the high school group: M1 and M2 fail to produce sufficient concentration in the 40-44 hours bin at the same time that they fail to generate enough individuals in the long hours region, and neither can generate the non-monotonicity found in the equilibrium wage profile.

[19]We again emphasize that this property is robust to the choice of θ_s within the interval $[1.4, 1.9]$. In particular, if we set $\theta_s = 1.9$ the implied values for θ_m are .328 and .451 for the bachelor and bachelor plus groups respectively, and so still imply a steep wage penalty for hours worked beyond 40.

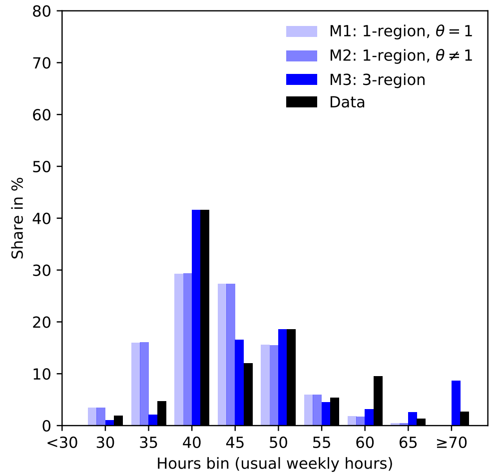
[20]As was true for our benchmark estimation, Online Appendix Figure D.2 shows that for each of these two groups we replicate fairly closely the standard deviation of log wages across the hours distribution, as well as mean hours and the standard deviation of hours across the wage distribution.

Figure 9: Comparing Fits of Models M1-M3: More than High School

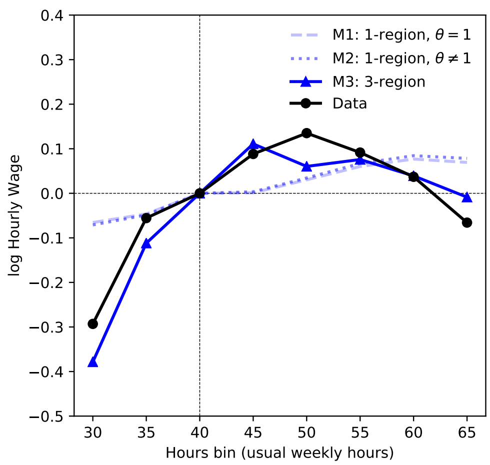
(a) Bachelor: Hours



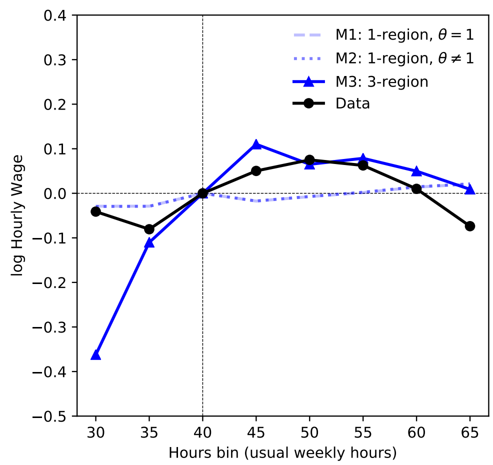
(b) Bachelor+: Hours



(c) Bachelor: Wages



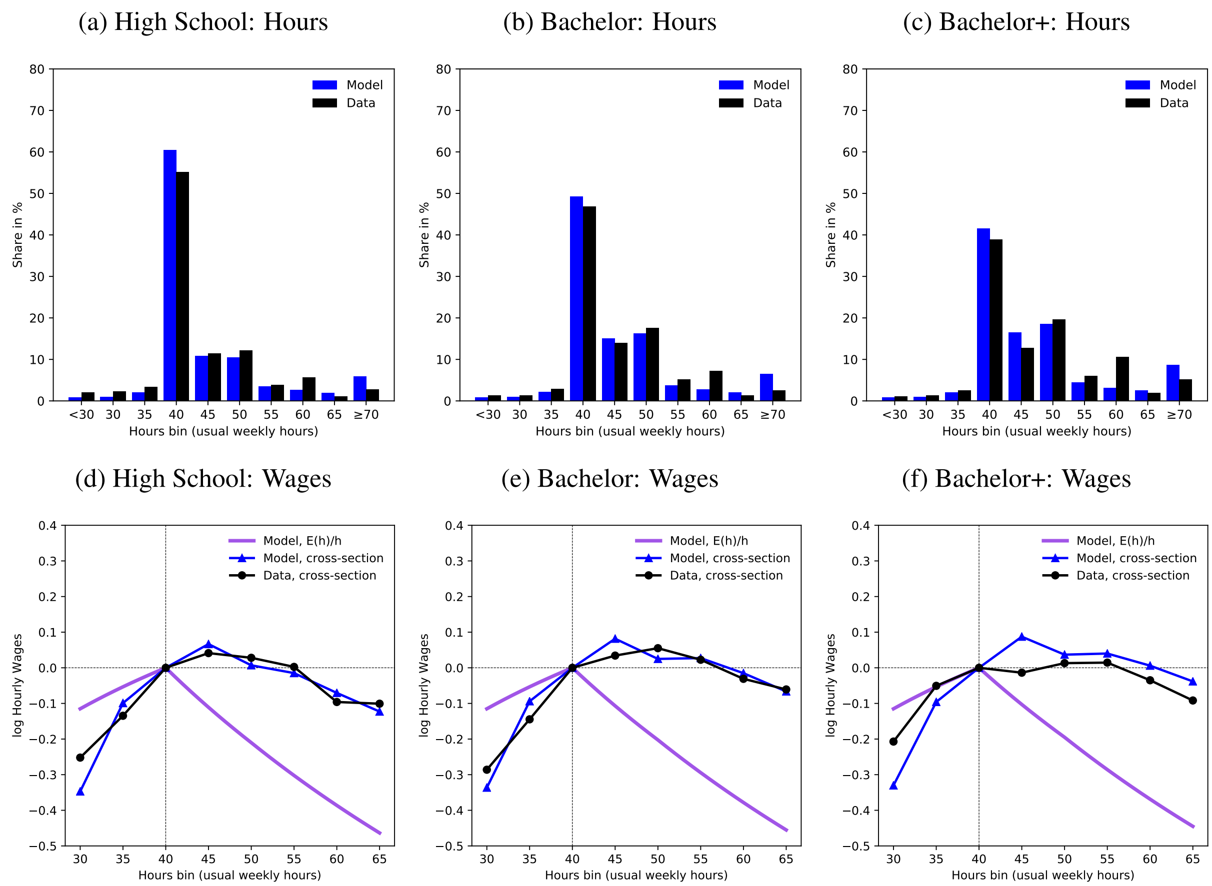
(d) Bachelor+: Wages



4.4 Implications for Other Age Groups

We chose to estimate our model on a sample of older workers because our model abstracted from human capital accumulation and we were mindful of the fact that future returns might be an important motive for some (especially younger) workers to work long hours. (See, e.g., the structural labor supply model of Imai and Keane (2004).) Having estimated our model on a sample of older individuals for whom future returns are arguably much less relevant, we now ask to what extent the forces that we estimate can account for the patterns observed for younger males. In particular, we focus on ages 30-34, as the results for nearby age groups are similar. In applying our estimated model to each of the three educational attainment groups we hold all parameters fixed with one exception: the standard deviation of σ_z is adjusted so as to match the standard deviation of wages. Allowing σ_z to change allows us to capture the well documented systematic

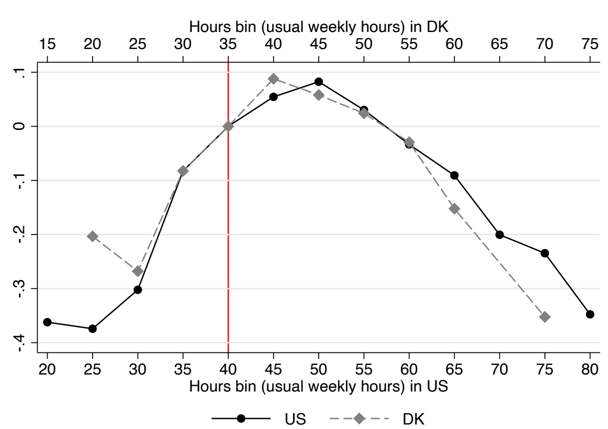
Figure 10: Fit of 3-Region Model (M3): 30-34 Year-Olds



increase in wage dispersion over the life cycle, consistent with either human capital accumulation (see, e.g., Huggett et al (2011)) or the accumulation of persistent productivity shocks (see, e.g., Guvenen (2009)). Importantly, we do not change the values of the θ_j or the extent of preference heterogeneity.

Figure 10 shows the results. The model accounts well for the features of the hours distribution, though it slightly underpredicts the fraction of individuals in the bachelor plus group that work 50 or more hours (36% in the model versus 44% in the data). The fit of the equilibrium wage profile is best for the high school and bachelor groups and a bit worse for the bachelor plus group. As noted earlier, there is no presumption that the static forces that our model estimates on older males will be sufficient to capture the cross-sectional patterns for younger workers. The fact that the model misses somewhat for the bachelor plus group is consistent with dynamic considerations being a more important factor for young individuals in this group. But overall, our estimated model also fits the data for younger workers reasonably well, suggesting that the static forces estimated using data on older workers also play an important role in shaping the cross-sectional patterns for a broader range of individuals.

Figure 11: Cross-Sectional Relationship between Wages and Hours in the US and Denmark



Note: The upper x-axis refers to Denmark and is shifted by 5 hours relative to the lower x-axis, which refers to the US. Data for Denmark are from the European Community Household Panel (ECHP), which ran from 1994-2001 and reports usual monthly earnings and usual weekly hours. To calculate weekly earnings, we assume 4.3 work weeks per month. We employ the same sample selection criteria as in our benchmark. Since the ECHP has no information on whether earnings and hours are imputed, we drop this criterion as well for the comparison sample for the CPS ORG covering the years 1994-2001. This has however virtually no impact on the wage-hours relationship for the US.

4.5 Interpreting the Kink in the Earnings Function

We view the kinked earnings function from our estimation procedure as reflecting two distinct but related forces. The first force reflects the extent to which average labor services (or efficiency units) per hour are affected by the length of the workweek. For example, if there are some set-up costs involved, then labor services may be convex in hours at low levels of hours, and if individuals become fatigued at long hours then there may be a concave region at higher levels of hours. Barzel (1973) and Rosen (1978) both emphasized this source of non-linearities. See Pencavel (2015) for a discussion of this issue and evidence in one particular setting.

The second force reflects coordination. Coordination issues exist both within and across production units. The assembly line is the classic example of a production process that requires workers within a given business to coordinate work schedules. But more generally, any business that has frequent interactions with other businesses has a desire to coordinate work hours with other businesses. The desire to coordinate will potentially affect how a firm values workweeks of different lengths. Yurdagul (2017) posits an aggregate production function in which inputs of different workers are complements, implying that workweeks for a particular worker are less valued as they move further from mean hours for other workers.

We view our estimated earnings function as reflecting both of these forces and do not attempt to separately identify them. We note that the two channels could be complementary; a tendency for a moderate increase in fatigue beginning around 40 hours might induce coordination around that point. But we close by offering one piece of evidence that we think highlights coordination as a factor. Differently than the US, Denmark has a modal workweek of 37 hours instead of 40 hours, and its concentration of workers in the 35-39 hours bin is comparable to that found in the US in the 40-44 hours bin (in both cases higher than 60%). Interestingly, Figure 11 shows that the wage-hours profile in Denmark looks very much like the wage-hours profile in the US except that it is shifted by one five hours bin. Consistent with the coordination narrative, this suggests that the peak of the wage-hours profile is influenced by the modal length of the workweek.

5 The Role of Other Income

Our benchmark specification assumes that individuals differ along only two dimensions: productivity and preferences. Textbook models of labor supply also stress income from other sources (e.g., spousal earnings, income from capital) as a factor that shapes the choice of hours worked. In the first part of this section we show that explicitly including dispersion in other income in our analysis has little effect on the estimated earnings function or the overall fit of the model. Our estimated model is also consistent with the results we reported in Section 2.5 regarding the low explanatory power of wealth and other income for usual hours worked.

In the second part of this section we highlight the fact that other income has significantly different effects on hours in our benchmark estimated model than in the estimated linear model M1. In our benchmark model, the response varies across the hours worked distribution, with individuals at or near a kink responding much less than individuals who are away from a kink. Because many individuals are located at a kink, this also implies a much smaller aggregate responses to changes in other income.

5.1 Adding Other Income to Our Benchmark Model

We extend our previous analysis to include an additional source of income, so that the budget equation for individual i now reads:

$$c_i = z_i E(h_i) + y_i$$

where $E(h_i)$ is the earnings function defined earlier and y_i is non-labor income, which will reflect all household income other than the individual's own labor earnings, e.g. spousal income or (implicit) capital income. An individual i will now be represented by the three vector (α_i, z_i, y_i). As before, we assume that (α_i, z_i) are joint lognormally distributed. We further assume that y_i is distributed lognormally with mean μ_y and standard deviation σ_y, and that it is uncorrelated with both α_i and z_i.[21]

We calibrate the value of σ_y to match the standard deviation of log non-labor income in the data, and we calibrate the value of μ_y so that the mean of log non-labor income relative to mean log labor earnings in the model is the same as in the data. We compute these moments using the SCF, since this is thought to be the highest quality data set in terms of information on wealth. We compute annual non-labor income as the sum of all sources of household income apart from male labor earnings plus 5% of total net wealth, and then convert this to a weekly value by dividing by 52 to align it with our measure of weekly earnings from the SCF. For males aged 50-54 with a high school education the resulting standard deviation of the log of our measure of non-labor income is equal to 1.38.[22]

Taking this distribution of other income as given, we now repeat our earlier estimation exercise for the subsample of males aged 50-54 with a high school education. A full set of results is contained in Online Appendix D.2; here we summarize the key findings. First, there is little change in the estimated parameters of the earnings function: the estimated values of θ_m and θ_l change from 0.058 and 0.034 in the benchmark estimation to 0.061 and 0.038 in the model with other income. Second, the overall fit of the model is essentially unaffected. Not surprisingly, the inclusion of an additional idiosyncratic factor does reduce the dispersion in tastes that are required, but this effect is modest: σ_α decreases from 1.61 to 1.50.

This exercise assumed that other income was uncorrelated with both α_i and z_i. To assess the empirical reasonableness of this assumption we compute the correlations of wages and hours with non-labor income in the model and the data. In the data these values are 0.091 and -0.090 respectively, and in our estimated model the implied values are 0.092 and -0.189. Both model implied values are close to their empirical counterparts, suggesting that the assumption of zero correlation is not unreasonable. The fact that our model generates a small correlation between hours and other income implies that it is also consistent with the results from Section 2.5, where we found that cross-sectional variation in other income had little explanatory power for cross-sectional variation in hours. In particular, using cross-sectional data generated from our estimated model to regress log usual hours on other income, we obtain an R^2 of 0.036, similar in magnitude to the values we found in Section 2.5.

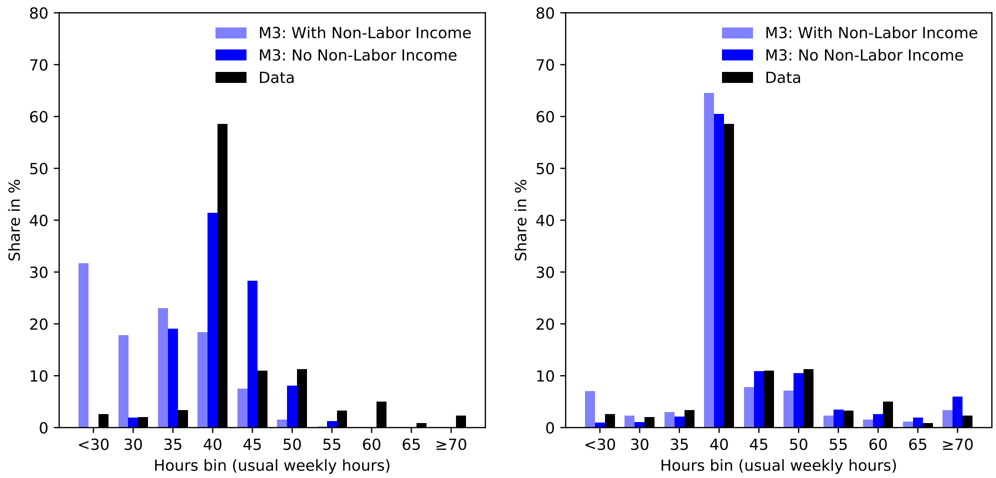
The results just described motivate our choice to adopt the more parsimonious specification without

[21]We assume that non-labor income is distributed lognormally as the overall wealth distribution is reasonably well approximated by this distribution apart form the upper tail. The top 1% is of second order importance for our analysis. First, many of these individuals are not employees and so are not part of our sample. Second, changing the behavior of a small number of individuals exerts a very small effect on the moments that we target.

[22]To calculate the standard deviation of the log of non-labor income we had to exclude observations that were negative. There were very few of these and excluding them had little impact on the mean and standard deviation of the unlogged data, so we do not think this is an issue for our calculations. The standard deviations for males aged 50-54 in the bachelor and bachelor plus groups were similar; 1.41 for the former and 1.30 for the latter.

Figure 12: Effect of Non-Labor Income on the Hours Distribution

(a) Linear Model (M1) (b) 3-Region Model (M3)



non-labor income as our benchmark.

5.2 Comparison With the Linear Model

The fact that adding dispersion in other income to our model has little effect on estimated parameter values and overall fit suggests that this dispersion has relatively small effects. In this section we examine this finding more closely and highlight that it is a distinctive feature of our model relative to both models M1 and M2. Because M1 and M2 are very similar in this regard, in the interests of space we confine ourselves here to a comparison of our benchmark model with the linear model M1.

The first exercise illustrates the very different impacts of other income on hours in the two models. To do this we take the estimated versions of M1 and M3 from Section 4, add the distribution of other income as measured above, and compute how this affects the distribution of hours worked. Figure 12 shows the results, with the left panel showing the effects in M1 and the right panel showing the results in M3. The figures have three bars for each five hours bin – one for the data, one for the benchmark estimated model without non-labor income and a third bar for the benchmark estimated model that adds non-labor income. A comparison of the two panels reveals that other income has dramatically larger effects on the hours distribution in M1 than in M3. In particular, whereas in M3 the mean of log hours changes from 3.774 to 3.742, in M1 it changes from 3.765 to 3.646. The change in mean log hours is more than three times as large in the linear model.[23] The reason for the very different results are the kinks in our estimated earnings function,

[23]The fact that other income has large effects in the linear model does not affect the ability of this model to fit the data. If we redo the estimation exercise when including other income, the fit is essentially unchanged. The estimated values of μ_α, σ_α and $\rho_{\alpha,z}$ change to largely undo the effects of other income.

Table 5: Impact of Receiving Other Income that is a Multiple k of Average Earnings

	Linear Model (M1)			3-Region Model (M3)		
$h(k=0)$	$\frac{h(k=.5)}{h(k=0)}$	$\frac{h(k=1)}{h(k=0)}$	$\frac{h(k=2)}{h(k=0)}$	$\frac{h(k=.5)}{h(k=0)}$	$\frac{h(k=1)}{h(k=0)}$	$\frac{h(k=2)}{h(k=0)}$
30	0.86	0.76	0.62	0.85	0.73	0.57
40	0.86	0.76	0.62	1.00	0.99	0.99
50	0.86	0.76	0.62	0.95	0.87	0.80
60	0.86	0.76	0.62	0.87	0.83	0.83

specifically the large kink in the earnings function at $h = 40$. Changes in non-labor income for people at the kink are relatively unlikely to cause them to change their desired hours. Because so many individuals are located at the kink, the aggregate effect tends to be small.

The above discussion is relevant for interpreting empirical evidence on labor supply responses. In particular, several recent studies have examined how individuals change labor supply in response to lottery winnings as a way to estimate the size of income effects, e.g. Cesarini et al (2017). Our framework implies that care must be taken when inferring preference parameters in this context. In particular, the mapping from preference parameters to the size of the implied income elasticity is highly non-linear, depending critically upon where the individual lies in the hours worked distribution.

To illustrate this we carry out a simple exercise in our estimated versions of Models M1 and M3. In particular, in each case we consider individuals at four different points in the hours distribution: 30, 40, 50 and 60, and ask how average log hours responds if they receive other income that is a multiple k of average earnings for individuals working the same number of hours. The rationale for these four values is that they correspond to different regions in the non-linear earnings specification: 30 hours corresponds to a region with convex earnings, 40 hours corresponds to a point with a large kink and a local maximum for wages/hour, 50 hours represents a point with a smaller kink and a decreasing, concave wage function, and 60 hours reflects a point without a kink but with a decreasing concave wage function.

Table 5 shows the results. While the effect of other income on hours in the linear model M1 is uniform across the hours distribution, it varies widely across the hours distribution in model M3. There is almost no response for individuals located at the large kink at 40 hours, even in the case of $k = 2.0$.[24] If we interpret the additional income as the perpetual flow from a one time wealth shock using a 5 percent value for annuitization, the case of $k = 2.0$ represents a wealth shock equal to 40 times earnings, i.e., a wealth shock of \$2 million for an individual earning \$50 thousand. The kink in the earnings function at 50 is much smaller than at 40, but the response of individuals located at this is still dampened significantly in M3

[24] The additional income does move a small mass of marginal individuals away from the kink, but this effect is very small.

relative to M1.

For individuals away from kinks, the response of hours to an exogenous increase in other income also varies with the slope of the earnings function. In particular, responses for an individual with $h = 30$ are modestly larger in M3 than M1. In contrast, responses for an individual with $h = 60$ are modestly smaller in M3 than M1 for $k = 0.5$ and 1.0.[25]

To summarize, our results suggest caution in using data from lottery winnings to estimate preference parameters underlying income elasticities. Through the lens of our model, individuals located at the kink are much less likely to adjust their hours, but this is driven by the shape of the earnings-hours relationship rather than primitive preferences. Because this effect is non-uniform across the hours distribution, it may have important implications for how government transfer programs affect hours along the intensive margin.

6 Implications for Labor Supply Responses to Wages and Taxes

A key feature of our estimated earnings function is a large kink at 40 hours, that generates both short and long hour wage penalties. In this section we carry out three exercises to highlight the consequences of our estimated earnings function for the response of labor supply to changes in wages. The first exercise embeds our non-linear earnings function into an otherwise standard heterogeneous agent-incomplete markets model to show that it significantly dampens the role of labor supply as a margin of adjustment in response to temporary wage shocks. Importantly, the extent of this dampening effect varies across the hours distribution and is largest for individuals located at a kink.

Motivated by this result, the other two exercises show that our model has important implications for mapping observed behavioral responses to changes in wages and taxes into underlying preference parameters. In particular, we first show that a standard procedure for estimating the Frisch elasticity delivers estimates that vary across the hours distribution in a way that is consistent with our model. Another strategy for estimating elasticity parameters in preferences is to rely on the extent of bunching observed at kinks created by the tax and transfer system. We show that our model with a kinked relationship between hours and earnings can dramatically reduce the amount of bunching that one should observe because of kinks in the tax and transfer system.

Because our analysis abstracts from any dynamic effects of hours on future earnings, these exercises should be understood as illustrating the impact of the static effects that we have estimated, and are meant to

[25]The effect of larger changes in other income for individuals with $h = 60$ in M3 are more complex. These individuals have the same value of h for both $k = 1$ and $k = 2$. This occurs because $k = 1$ pushes these individuals to the kink at 50 and the increase of k from 1.0 to 2.0 is just marginally sufficient to start moving them away from the kink but not sufficient to create a change to two decimal points.

complement the analysis in Imai and Keane (2004) regarding how dynamic effects matter for labor supply responses. We think that these are useful and important exercises precisely because this static component has not previously been studied.

6.1 Labor Supply in Incomplete Markets Models

We consider an infinitely lived individual with preferences given by:

$$\sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1-(1/\sigma)} c_t^{1-\frac{1}{\sigma}} - \frac{\alpha}{1+(1/\gamma)} h_t^{1+\frac{1}{\gamma}} \right],$$

who faces the following period budget constraint:

$$c_t + a_{t+1} = z_t E(h_t) w + (1+r) a_t,$$

where z_t is idiosyncratic productivity, $E(h)$ is our estimated earnings function, w is the (constant) wage per efficiency unit of labor services, a_t is assets and r is the (constant) interest rate. We assume that assets must be non-negative: $a_t \geq 0$. Idiosyncratic productivity is stochastic and follows the process:

$$\log z_{t+1} = \rho_z \log z_t + \varepsilon_t$$

where ε_t is normally distributed with mean μ_ε and standard deviation σ_ε.[26]

To assess the importance of our estimated earnings function for the role of labor supply in this setting we compare three specifications. The first assumes that hours are exogenously fixed, so that labor income follows the same process as z, and the individual uses savings to smooth consumption. The second specification follows Pijoan-Mas (2006), assuming $E(h) = h$ and that the individual can freely adjust hours. In this setting, the individual can reallocate hours of work from periods of low productivity to periods of high productivity and so generate higher earnings per hour.

The third specification replaces the assumption of $E(h) = h$ with our benchmark estimates for males aged 50-54 with a high school degree. Intuitively, the role of variable labor supply in this context is likely to depend on where an individual lies in the hours distribution and so in what follows we will consider the same four settings for average hours as in our earlier analysis of income effects: 30, 40, 50, and 60.

To proceed with the comparison we normalize w to unity and consider the following parameterization,

[26] We assume a simple AR(1) process to maintain comparability with Pijoan-Mas (2006). In Online Appendix E we find similar results for a stochastic process with both persistent and transitory shocks.

Table 6: Effects of Endogenizing Hours

	mean h	std h	CEV
linear earnings	40	0.24	3.5%
non-linear earnings	40	0.01	0.0%
non-linear earnings	30	0.27	6.1%
non-linear earnings	50	0.10	0.1%
non-linear earnings	60	0.14	0.1%

drawn from Pijoan-Mas (2006): $\sigma = 0.69$, $\gamma = 0.50$, $\rho_z = 0.92$, $\sigma_z = 0.20$, $\beta = 0.94$, and $r = 0.05$. Interpreting a period to be a year these parameters are all quite standard. We approximate the AR(1) process for $\log z$ using a Tauchen procedure with seven grid points and solve for the ergodic distribution that characterizes the behavior of this individual. Equivalently, this is the stationary distribution for an economy consisting of a large number of individuals that each solve this same maximization problem with the shocks iid across individuals. We choose the value of α so as to target average hours in the ergodic distribution and consider the four targets mentioned earlier: 30, 40, 50, and 60. Note that the values of α will differ across the linear and non-linear specifications.

We compute the effect of allowing for endogenous labor supply responses starting from the specification in which we restrict the individual to have constant hours equal to the mean in the ergodic distribution. When doing this calculation we solve for the resulting transition from the initial ergodic distribution to the new ergodic distribution and compute two statistics: the variation in log hours in the final ergodic distribution and the welfare gain measured in consumption equivalent variation, including the transition path. The welfare gain will necessarily be positive since the individual can always choose fixed hours and so replicate the original allocation.

Table 6 shows the results. With linear earnings the value of mean hours does not affect the results given that our preferences feature constant elasticities, so we only report results for the $h = 40$ case. The first row confirms one of the messages in Pijoan-Mas (2006) – given the opportunity to vary labor supply we see that hours vary substantially, and this produces a significant increase in welfare. In contrast, the second row shows that these results effectively disappear if we consider the non-linear earnings specification and an individual who on average is located at the kink of the earnings function at 40. That is, the variation in hours nearly disappears, and not surprisingly, so do the welfare gains. The fact that the variation in hours is so small implies that the individual is not just at the kink on average but is almost always at the kink.

The final three rows show that the impact of endogenous hours is quite different if we consider an individual away from the kink. Interestingly, these impacts differ depending on which direction we move away from the kink at 40. Consider first the case of an individual who on average works 30 hours. In this situation, the impact of allowing for variable hours is larger than in the linear case. In this region, working more hours when productivity is high is even more powerful in terms of generating additional income given that the individual is in a region where earnings are convex in hours. As a result we see that hours are somewhat more volatile than in the linear earnings case, and that the welfare improvement is more than one and a half times larger. Consider next the case of an individual that works 50 hours on average. While also located at a kink on average, this kink is much smaller than the kink at 40, and the wage function does not have a local maximum at 50. This individual displays a variation in hours that is intermediate between those for the individuals at 30 and 40 hours, but the welfare gain is almost negligible; moving hours to periods of high productivity is no longer very powerful in terms of generating higher income per hour. Finally, consider the individual with mean hours of 60. Because this individual is not at a kink, movements in productivity will generate larger movement in hours than for the individual at 50 hours, but once again the welfare benefit associated with the ability to adjust hours is negligible.

The key message from this exercise is that inserting our non-linear earnings technology into an otherwise standard incomplete markets economy has first order implications for the role of labor supply, but that these effects are very non-uniform across the hours distribution. While the majority of individuals in reality are at the kink at 40 hours and so might well be approximated by an indivisible labor model, this approximation does not apply to the still significant mass of individuals who are away from the kink.

A simple implication of the preceding analysis is that individuals working 40 hours are much less likely to change their hours of work from one year to the next. This implication is strongly supported by the data. Using the panel component of the CPS-ORG sample we can examine the propensity for individuals to change their usual hours of work depending upon their initial level of hours. For this exercise we consider three regions for usual weekly hours: those between 10 and 35 (short hours), those between 35 and 49 (middle hours) and those with at least 50 hours (long hours), and compare changes in usual weekly hours at interviews one year apart (the 4th and 8th interviews).

The values reported below are for males aged 25-64 with education of high school and above; the patterns are very similar if we disaggregate by age and education. We first compute the percentage of individuals who report a change in usual hours across the interviews. Consistent with our theory, the three values are 63.0, 27.4 and 64.6 for the short, middle and long hours regions respectively. For individuals who report exactly 40 hours the percentage that change is 16.3. One concern with this calculation is that it may simply reflect a larger role for measurement error for individuals who do not report 40 hours. To partially address this we carry out a second calculation in which we compute the probability of usual weekly hours changing by more than plus or minus five hours between interviews. These percentages are 34.1, 9.0, and 41.6 for the short, middle and long hours regions respectively. Put differently, individuals in the short and long hours regions are more than three times as likely to change their usual hours by more than plus or minus five hours than are individuals in the middle hours region. While this descriptive evidence is consistent with our model, it does not allow us to conclude that individuals at different points of the hours distribution respond differently to a given shock. Examining this issue further requires a more structural approach, which we pursue in the next subsection.

6.2 Estimating Frisch Elasticities from Panel Data

A large literature has used life cycle variation in hours and wages to estimate the individual elasticity of intertemporal substitution (IES). (See for example, MaCurdy (1981), Browning et al (1985) and Altonji (1986).) The standard assumption in these analyses is that earnings are linear in hours worked. The discussion in the previous subsection highlights why the non-linearities that we estimate might also affect the interpretation of these estimation exercises; individuals located at the kink in the earnings function require very large changes in wages in order to generate changes in hours, and this will tend to dampen the estimated IES for these individuals. But this dampening reflects reduced opportunities for intertemporal substitution and not necessarily reduced willingness to do so.

In the interest of space, we move immediately to the key estimating equation in this literature, and refer the interested reader to one of the previously mentioned papers for details on its derivation. Specifically, consider the following specification to be run on panel data:

$$\Delta \log h_{it} = a + b \Delta \log w_{it},$$

where a and b are constants. The key result from this literature is that under a set of standard assumptions, the coefficient b will be an estimate of the preference parameter that governs intertemporal substitution. In this subsection we explore the extent to which this estimate varies across the hours distribution. To do this we implement a standard estimation exercise using the data (PSID 1972-1997) and codes from Bredemeier et al (2019). We refer the reader to their paper for more details. We restrict their sample to males with at at least four first differences observations such that we have at least five observations to construct average hours. Our sample is thus a bit smaller than the original sample in Bredemeier et al (2019).

In addition, we sort workers into weekly hours bins based on their average annual hours worked divided by 52 (to the extent that most males work full year this assumption is not unreasonable) into three bins:

Table 7: IES Estimates

	Altonji (1986)				Bredemeier et al (2019)			
	All	< 35	$37-43$	> 50	*All*	< 35	$37-43$	> 50
IES	0.30	0.47	0.27	0.88	0.52	0.66	0.51	1.07
	(0.11)	(0.67)	(0.15)	(0.30)	(0.23)	(1.35)	(0.32)	(0.64)
Observations	12043	553	5417	1845	12043	553	5417	1845
Individuals	1415	71	598	235	1415	71	598	235

Note: Sample is based on the PSID for the years 1972-1997. For details of sample construction and regression specifications, see Bredemeier et al (2019). We in addition restrict the sample to men with at least four first differences observations.

below 35, between 37 and 43, or at least 50. In our estimated model it is not the case that everyone in the 37-43 hours bin is at the kink at 40 hours. However, the fraction of individuals in the 37-43 hours range that are at the kink is 87 percent such that this group is dominated by individuals at the kink. Our estimated model also included a kink at 50 hours that could be relevant for the group with average hours of 50 or more. But in our estimated model, only 21 percent of individuals with $h \geq 50$ are at a kink, so the kink is far less important for this group. It follows that our estimated model suggests that standard Frisch elasticity estimates should be lower for those in the 37-43 hours range.

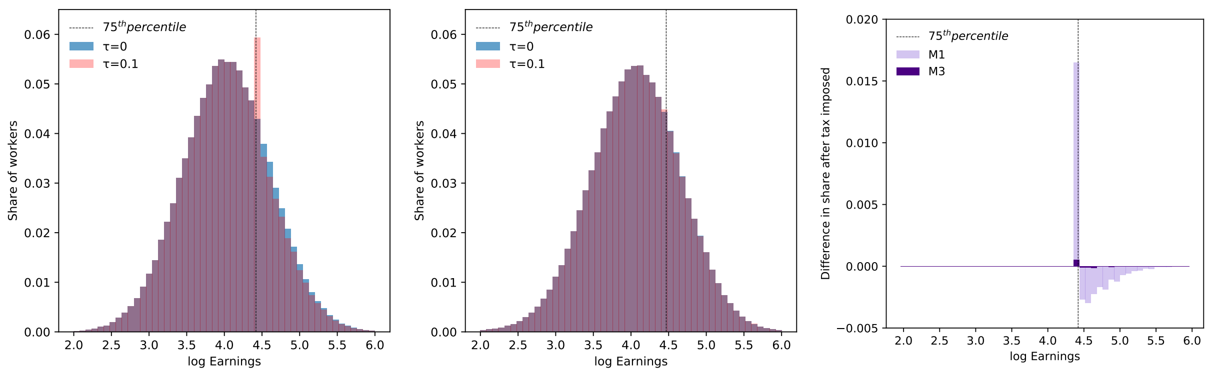
A key contribution of Bredemeier et al (2019) is to develop an estimation procedure that generalizes Altonji (1986) to allow for borrowing constraints. Consistent with the work of Domeij and Floden (2006), they find that borrowing constraints do lead to a substantial downward bias in the estimated IES and which can be seen when comparing the two columns labeled "All" in Table 7. The result of interest for us is however the comparison across the different hours bins. The key finding is that whether we use the Altonji (1986) specification or the Bredemeier et al (2019) specification we find a U-shaped pattern for the estimates of the IES across the hours worked distribution, with the lowest value for those working in the 37-43 hours interval. While standard errors preclude definitive conclusions about how estimates vary across the three regions, this pattern is consistent with the prediction that the responsiveness to shocks varies across the hours distribution.

6.3 Estimating Elasticities from Bunching at Tax Kinks

Our results also have potential implications for the literature that has used bunching to produce estimates of taxable income elasticities. (An early contribution is Saez (2010); see Blomquist (2021) for a recent assessment of this method and a more complete set of references.) Loosely speaking, this method leverages the fact that kinks in the tax system create incentives for individuals to bunch in terms of earnings, and that

Figure 13: The Effect of Introducing a Kinked Earnings Tax on the Earnings Distribution

(a) Linear Model (M1) (b) 3-Region Model (M3) (c) Change in the Distribution



the extent of bunching reveals information about labor supply elasticities; if there is little bunching at tax kinks then the relevant labor supply elasticity must be small. But the large kink that we estimate in the earnings function creates an incentive for individuals to bunch in terms of hours. If the kink in the earnings function E is large relative to the kink generated by the tax system there may be little bunching at the tax kinks precisely because of the incentives to bunch at 40 hours.

To illustrate the potential significance of this issue we introduce a kinked tax function into the estimated versions of the linear model M1 and our benchmark model M3 and compare the extent of bunching in the two models. In particular, we create a kink in the tax schedule at the 75th percentile of earnings by assuming a tax rate τ equal to 0.10 for all income in excess of this level of earnings. Figure 13 contrasts the effects in the two models. Panel (a) and (b) illustrate how the kinked tax function affects bunching with regard to earnings in the two models. In each of these panels the purple region shows the area of overlap for the cases of $\tau = 0$ and $\tau = 0.10$, the blue bars fill in the remaining mass for $\tau = 0$, and the pink bars fill in the remaining mass for $\tau = 0.10$. Panel (a) clearly shows that in the linear model the kink in the tax function moves a significant amount of mass from above the kink and concentrates it at the kink. While Panel (b) shows similar qualitative effects, the magnitude is much less and the reductions in mass above the kink are essentially not visible. To highlight the different size of the effects, Panel (c) shows the changes in the distributions of individuals across earnings levels between $\tau = 0$ and $\tau = 0.10$ for the two models: the extent of bunching is larger by an order of magnitude between M1 and M3.

Varying the position of the kink in the earnings distribution in the above example has little effect on the key message. The reason for this is that most regions within the male earnings distribution are dominated by individuals working 40 hours. The main message from this example is that it is important to incorporate information about usual weekly hours and their determinants when mapping the extent of bunching into an elasticity. Importantly, this mapping may be quite different when looking at subsamples with different propensities for part-time work, and may explain why more bunching is observed in some settings than others.

7 Non-Monetary Compensation

Our empirical and theoretical analysis has focused on the relationship between current hours of work and current monetary compensation.[27] One potential concern is that part of the return for working long hours might take the form of non-monetary compensation that reflects better working conditions, such as a more prestigious job title, greater "power", more administrative support, a better office, or a closer parking spot. If this is the case, our measures of the monetary compensation for workers working longer hours might present a misleading picture of the total benefits of working longer hours.

While the data does not allow us to address this issue directly, we can explore it indirectly in the context of our structural model. Specifically, we extend our model to assume an additional form of compensation for hours worked beyond 40. We model this as a latent variable, and so define the *effective* consumption of individual i, denoted by $\hat{c}_i$ as:

$$\hat{c}_i = z_i \cdot (E(h_i) + \delta W(40) \max\{h_i - 40, 0\})$$

where $E(h_i)$ is the earnings function that we defined previously. For comparison with the earlier exercises, we will refer to this as Model 4 (M4). While we motivated our analysis by noting the possibility of non-monetary components of current compensation, we note that this specification can also be interpreted as a reduced form way of capturing future benefits. Note that δ measures the magnitude of this effect relative to monetary earnings when working 40 hours.

Before proceeding we introduce some additional concepts and notation to reflect the different notions of earnings in this extension. To maintain consistency with our previous definitions, we will continue to let $E(h)$ and $W(h)$ denote monetary labor earnings and monetary wages. But we introduce the two functions $\hat{E}(h)$ and $\hat{W}(h)$ to reflect *effective* earnings and *effective* wages, i.e.,

$$\begin{aligned}
\hat{E}(h) &= (E(h) + \delta W(40) \max\{h - 40, 0\}) \\
\hat{W}(h) &= \hat{E}(h)/h = W(h) + \delta W(40) \frac{\max\{h - 40, 0\}}{h}.
\end{aligned}$$

[27] As noted earlier in our empirical analysis, our main results focus purely on labor earnings, but we found that the patterns did not change when we include information about other components of compensation such as health insurance and employer sponsored retirement programs.

Figure 14: Fit of Model with and Without Nonpecuniary Returns to Long Hours

(a) Hours Distribution

(b) Wage-Hours Profile

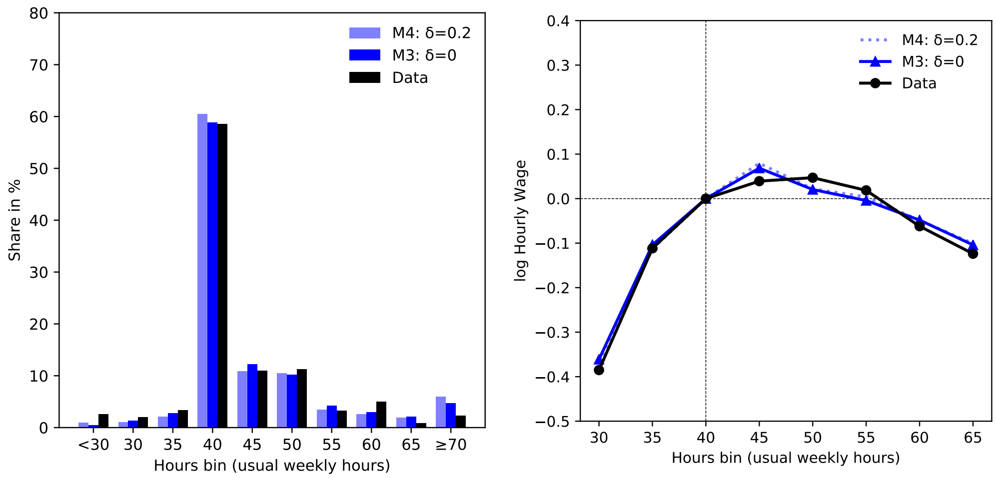


Table 8: Estimated Parameters, Males 50-54 With High School Degree

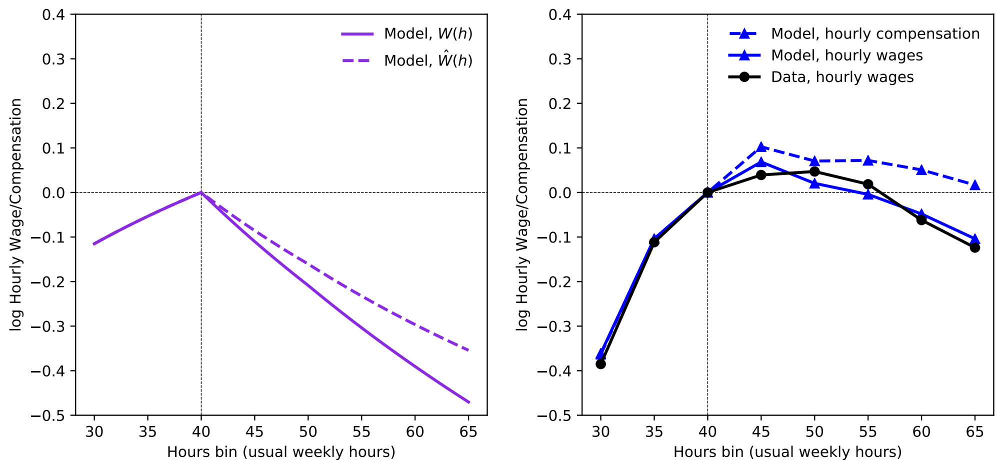
	δ	μ_α	σ_α	σ_z	$\rho_{\alpha,z}$	σ_m	θ_s	θ_m	θ_l
M3	0.00	-13.64	1.61	0.589	-0.33	0.04	1.40	0.058	0.034
M4	0.20	-12.32	0.96	0.589	-0.32	0.04	1.40	0.065	0.002

We now repeat our earlier estimation exercise for males aged 50-54 with a high school degree, including δ as an additional variable to be estimated, holding only σ_m fixed at its previously estimated value. Figure 14 compares how our estimated models M3 and M4 compare in terms of fitting moments of the hours distributions and the equilibrium wage profile. Model M3, with $\delta = 0$, already offers a good fit to the data, and the main takeaway from this figure is that freeing up δ offers little in terms of improving this fit. However, introducing δ could impact the estimated parameters and/or the implied interpretation of the patterns in the data, so we turn to this next.

Table 8 displays the estimated parameters for models M3 and M4. The estimated value for δ is 0.20. This estimate implies that an individual working 50 hours will receive five percent of their overall earnings in the form of the non-monetary earnings; the corresponding figure for an individual working 60 hours is nine percent. The largest change in parameter values is the decrease in the value of σ_α. This change is intuitive; if we add an additional incentive for individuals to work long hours then heterogeneous tastes for work become less important in inducing individuals to work long hours and so we require less dispersion in tastes for work. Less dispersion in tastes for work in turn requires a slightly smaller kink at 40 to achieve the large concentration of individuals in the 40-44 hours bin. There is a small decrease in θ_l, which partially

Figure 15: Wages vs. Total Compensation in Model with Nonpecuniary Returns (M4)

(a) Eq. Wage and Compensation Functions (b) Wage and Compensation Profiles



offsets the additional incentives to work longer hours. But note that there are limits to the extent to which θ_m and θ_l can be adjusted to offset a higher value of δ since we are still targeting moments for monetary compensation.

Introducing $\delta > 0$ does not affect the key properties of our estimated earnings functions, as determined by the θ_j. That is, we continue to find significant monetary wage penalties for long hours. However, in the model with $\delta > 0$, the object of interest from a worker's perspective is *effective* wages and not simply monetary wages. For this reason we are most interested in the implied properties of $\hat{W}(h)$. Panel (a) of Figure 15 shows the two estimated functions $W(h)$ and $\hat{W}(h)$. The key message from this figure is that the penalty for long hours remains significant even when considering effective wages.

It is also of interest to examine the effect of δ on the equilibrium profile for effective wages per hour and in particular the extent to which it differs from the equilibrium profile for monetary wages per hour. (Note that while we observe the equilibrium profile for monetary earnings in the data, we do not observe the equilibrium profile for effective earnings.) Panel (b) of Figure 15 shows these results. Whereas the equilibrium monetary wage profile is decreasing beyond 50 hours, the equilibrium effective wage profile is relatively flat. But importantly, the flatness of the equilibrium effective wage profile does not imply that there is not an effective wage penalty for working long hours; as panel (a) showed, the effective wage penalty for working long hours remains very substantial.

8 Conclusion

This paper focuses on understanding the economic forces that shape the cross-sectional distribution of hours and wages and the implications of these forces for labor supply responses. Two empirical observations play a key role in our analysis. First, although the distribution of usual weekly hours features a large concentration of individuals who work around 40 hours, more than 20 percent of men work 50 or more hours. Second, the profile of mean wages versus usual hours is non-monotonic–increasing below 50 hours and decreasing above 50 hours.

We argue that simple textbook models of labor supply cannot account for these facts jointly. This motivates us to extend these models to feature a non-linear earnings function. Our estimated model matches closely the two key features of the cross-sectional data that we highlight. Strikingly, we estimate not only a significant part-time wage penalty but also a large penalty for workers who choose to work more than 40 hours. We estimate that selection on unobservables plays an important role in shaping the cross-sectional profile of wages across the hours distribution, implying that the menu of hours and wages available to workers cannot be directly inferred from the cross-section data.

Our estimated earnings function has important implications for how labor supply responds to changes in wages, tax rates and other income. Individual responses may vary significantly across the hours distribution, and aggregate responses will be dampened due to the fact that individuals located at a kink in the earnings function are less likely to respond. In particular, responses for individuals who tend to work more than 40 hours will be larger than the average response. This finding may have important implications for optimal tax rates on top earners since this group tends to have usual hours above 40.

We close by emphasizing some important directions for extensions and future work. As noted several times, our analysis has focused entirely on static effects. For this reason, we estimate our model on a sample of older males for whom dynamic considerations such as human capital accumulation are less likely to be important. A priority for future work is to extend the analysis to a dynamic setting in which we can assess both the static and dynamic dimensions of earnings functions. An important implication of our analysis is that ignoring static effects could produce large biases in estimating the dynamic effect of current long hours on future wages: our estimates imply that an individual who works 60 hours this year and 40 next year will experience a wage increase of almost ten log points purely from static effects.

It is also of interest to extend our estimation to females. As noted earlier, this will require introducing an extensive margin into the analysis in order to capture the greater significance of the participation margin for females. Our estimates for males could have interesting implications for variation in labor supply responses by gender. In particular, we found that responses to temporary wage changes were much greater for part-time workers than for other workers. Since many more females than males work part-time, this would suggest that estimated labor supply elasticities might differ between males and females, which is indeed the case in the data (see, e.g., Altonji (1986) and Bredmeier et al (2019)). But importantly, it would be of interest to disentangle the roles of gender and part-time work status.

References

Aaronson, D. and E. French. 2004. "The Effect of Part Time Work on Wages: Evidence from Social Security Rules." Journal of Labor Economics 22, 329-352.

Abowd.J. and D. Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." Econometrica 57, 411-445.

Altonji, J., 1986. "Intertemporal Substitution in Labor Supply: Evidence from Micro Data." Journal of Political Economy 94, S176-S215.

Altonji, J., and C. Paxson. 1988. "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-offs," Journal of Labor Economics, 6, 254–276.

Ameriks, J., J. Briggs, A. Caplin, M. Lee, M. D. Shapiro and C. Tonetti. 2020. "Older Americans Would Work Longer If Jobs Were Flexible." American Economic Journal: Macroeconomics 12, 174-209.

Badel, A., M. Huggett, and W. Luo. 2020. "Taxing Top Earners: A Human Capital Perspective." The Economic Journal, 130(629), 1200-1225.

Barlevy, G. and D. Neal. 2019. "Allocating Effort and Talent in Professional Labor Markets." Journal of Labor Economics 37, 187-246.

Barzel, Y. 1973. "The Determination of Daily Hours and Wages." Quarterly Journal of Economics 87, 220-238.

Bredemeier, C., J. Gravert, and F. Juessen. 2019. "Estimating Labor Supply Elasticities with Joint Borrowing Constraints of Couples." Journal of Labor Economics 37, 1215-1265.

Blomquist, S., W. Newey, A. Kumar and C. Liang. 2021. "On Bunching and Identification of the Taxable Income Elasticity." Forthcoming, Journal of Political Economy.

Boerma, J., and L. Karabarbounis. 2021. "Inferring Inequality With Home Production." Forthcoming, Econometrica.

Borjas, G. 1980. "The Relationship Between Wages and Weekly Hours of Work: The Role of Division Bias."Journal of Human Resources, 15, 409-423.

Browning, M., A. Deaton, and M. Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands Over the Life Cycle." Econometrica 53, 503-544.

Cesarini, D., E. Lindqvist, M. Notowidigdo, and R. Ostling. 2017. "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries." American Economic Review 107, 3917-3946.

Cha, Y. and K. Weeden. 2014. "Overwork and the Slow Convergence in the Gender Gap in Wages." American Sociological Review 79, 457-484.

Chang, Y., S. Kim, K. Kwon, and R. Rogerson. 2019. "Individual and Aggregate Labor Supply in Heterogeneous Agent Economies with Intensive and Extensive Margins." International Economic Review, 60, 1-21.

Chang, Y., S. Kim, K. Kwon, and R. Rogerson. 2020. "Cross-Sectional and Aggregate Labor Supply." European Economic Review, 126, 103457.

Cogan, John. 1981. "Fixed Costs and Labor Supply." Econometrica 49, 945-963.

Cortes, P., and Pan, J. 2019. "When time binds: Substitutes for household production, returns to working long hours, and the skilled gender wage gap. " Journal of Labor Economics, 37(2), 351-398.

Denning, J., L. Lefgren, B. Jacob and C. vom Lehn. 2019. "The Return to Hours Worked Within and Across Occupations: The Implications for the Gender Wage Gap." NBER Working Paper #25739.

Domeij, D. and C. M. Floden. 2006. "The Labor Supply Elasticity and Borrowing Constraints: Why Estimates are Biased." Review of Economic Dynamics 9, 242-262.

Erosa, A., L. Fuster and G. Kambourov. 2016. "Towards a Micro-Founded Theory of Aggregate Labor Supply." Review of Economic Studies 83, 1001-1039.

Flood, S., M. King, R. Rodgers, S. Ruggles, and J. Warren. 2020 "Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]. " Minneapolis, MN: IPUMS.

Frazis, H. and J. Stewart. 2014. "Is the Workweek Really Overestimated?" Monthly Labor Review, U.S. Bureau of Labor Statistics, June. https://doi.org/10.21916/mlr.2014.21.

French, E. 2005. "The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behavior." Review of Economic Studies 72, 395-427.

Fuentes, J., and E. Leamer. 2019. "Effort: The Unrecognized Contributor to US Income Inequality." NBER Working Paper #26421.

Gicheva, D. 2013. "Working Long Hours and Early Career Outcomes in the High-End Labor Market." Journal of Labor Economics 31, 785-824.

Gicheva, D. 2020. "Occupational Social Value and Returns to Long Hours." Economica, 87(347), 682-712.

Goldin, C. 2014. "A Grand Gender Convergence: Its Last Chapter." American Economic Review 104, 1091-1119.

Guvenen, F., 2009. "An Empirical Investigation of Labor Income Processes." Review of Economic Dynamics 12, 58-79.

Heathcote, J., K. Storesletten, and G. L. Violante. 2010. "The Macroeconomic Implications of Rising Wage Inequality in the United States," Journal of Political Economy 118, 681-722.

Heathcote, J., K. Storesletten, and G. L. Violante. 2014. "Consumption and Labor Supply With Partial Insurance: An Analytical Framework," American Economic Review 104, 2075–2126.

Heckman, J., L. Lochner and C. Taber. 1998. "Explaining Rising Wage Inequality: Explorations With a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." Review of Economic Dynamics 1, 1-58.

Hirsch, B. 2005. "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills." Industrial and Labor Relations Review 58, 525-551.

Hofferth, S., S. Flood, M. Sobeck and D. Backman 2020. "American Time Use Survey Data Extract Builder: Version 2.8 [dataset]." College Park, MD: University of Maryland and Minneapolis, MN: IPUMS.

Huggett, M., Ventura, G., and Yaron, A. 2011. "Sources of lifetime inequality. " American Economic Review, 101(7), 2923-54.

Keane, M. 2011. "Labor Supply and Taxes: A Survey." Journal of Economic Literature 49, 961-1075.

Kuhn, P. and F. Lozano. 2008. "The Expanding Workweek? Understanding Trends in Long Work Hours Among Men." Journal of Labor Economics 26, 311-343.

Imai, S. and M. P. Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation." International Economic Review 45, 601-641.

Krueger, D., F. Perri, and L. Pistaferri. 2010. “Cross-sectional Facts for Macroeconomists.” Review of Economic Dynamics 13, 1-14.

MaCurdy, T. 1981. “An Empirical Model of Labor Supply in a Life Cycle Setting.” Journal of Political Economy 89, 1059-1085.

Michelacci, C., and J. Pijoan-Mas. 2012. “Intertemporal Labor Supply With Search Frictions.” Review of Econonic Studies 79, 899-931.

Moffitt, R. 1984. “The Estimation of a Joint Wage-Hours Labor Supply Model.” Journal of Labor Economics 2, 550-566.

Pencavel, J. 2015. “The Productivity of Working Hours.” The Economic Journal, 125 (589), 2052-20776.

Pijoan-Mas, Josep. 2006. “Precautionary Savings or Working Longer Hours.” Review of Economic Dynamics 9, 326-352.

Robinson, J., S. Martin, I. Glorieux and J. Minnen. 2011. “The Overestimated Workweek Revisited.” Monthly Labor Review, June, 43-53.

Rogerson, R., and J. Wallenius. 2009. “Micro and Macro Elasticities in a Life Cycle Model with Taxes.” Journal of Economic Theory 144, 2277-2293.

Rosen, H. 1976. “Taxes in a Labor Supply Model with Joint Wage-Hours Determination.” Econometrica 44, 485-507.

Rosen, S. 1978. “The Supply of Work Schedules and Employment.” in Work Time and Employment (Washington, DC: National Commission for Manpower Policy).

Ruggles, S., S. Flood, R. Goeken, J. Pacas, M. Schouweiler and M. Sobek. 2021. “IPUMS USA: Version 11.0 [dataset].” Minneapolis, MN: IPUMS.

Saez, E. 2010. “Do Taxpayers Bunch at Kink Points?” American Economic Journal : Economic Policy 2, 180-212.

Shao, L., F. Sohail and E. Yurdagul. 2021. “Labor Supply and Establishment Size.” Working Paper, Universidad Carlos III.

Trejo, S. 1991. “The Effects of Overtime Pay Regulation on Worker Compensation.” American Economic Review 81, 719-740.

Weeden, K., Y. Cha and M. Bucca. 2016. “Long Work Hours, Part-Time Work, and Trends in the Gender Gap in Pay, the Motherhood Wage Penalty, and the Fatherhood Wage Premium.” Russell Sage Foundation Journal of the Social Sciences 4, 71-102.

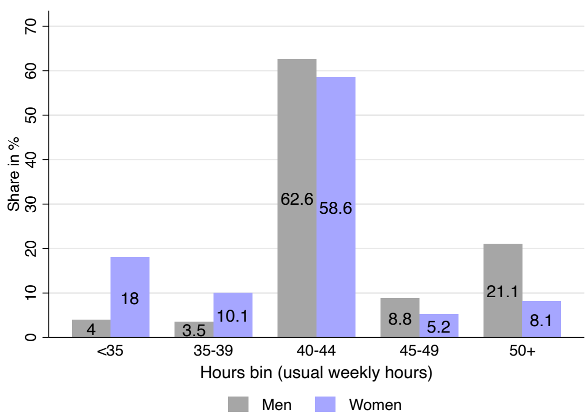
Yurdagul, E. 2017. “Production Complementarities and Flexibility in a Model of Entrepreneurship.” Journal of Monetary Economics 86, 36-51.

A Appendix Figures and Tables

A.1 Figures

Figure A.1: Main Facts by Gender

(a) Distribution of Usual Weekly Hours



(b) Long Hours Distribution

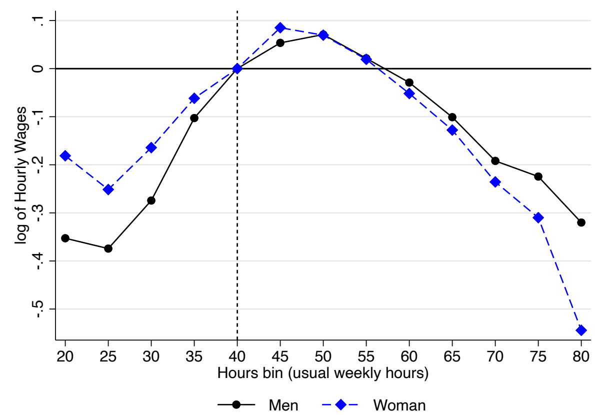


Figure A.2: Time-Series in CPS ORG

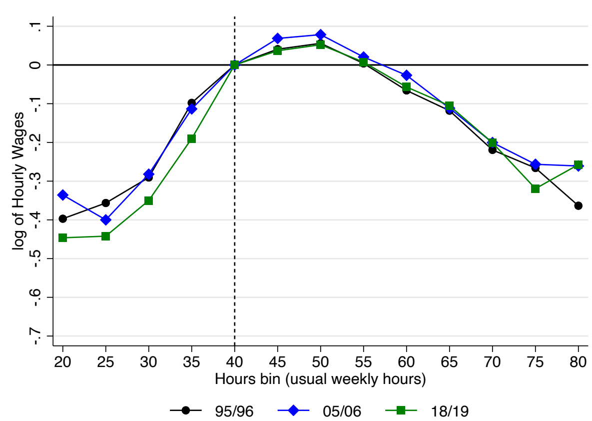
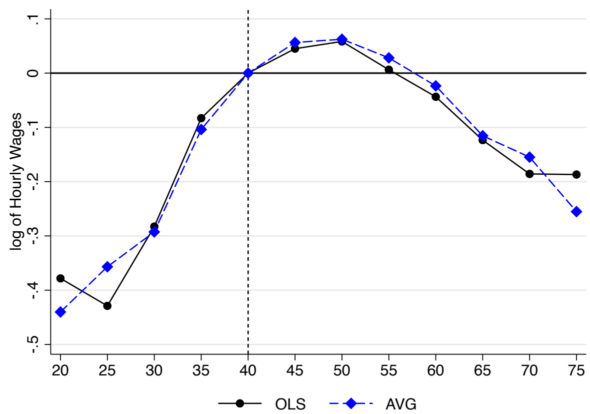


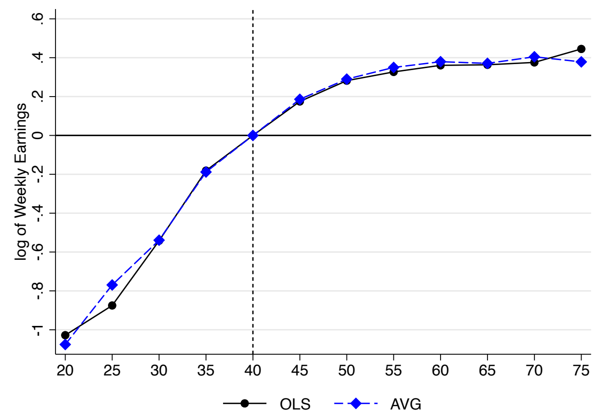
Figure A.3: Cross-Section vs. Cross-Section of Individual Averages

CPS ORG

(a) log Weekly Wage

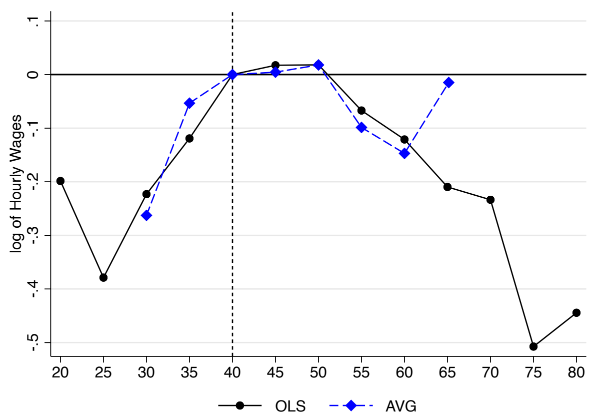


(b) log Weekly Earnings

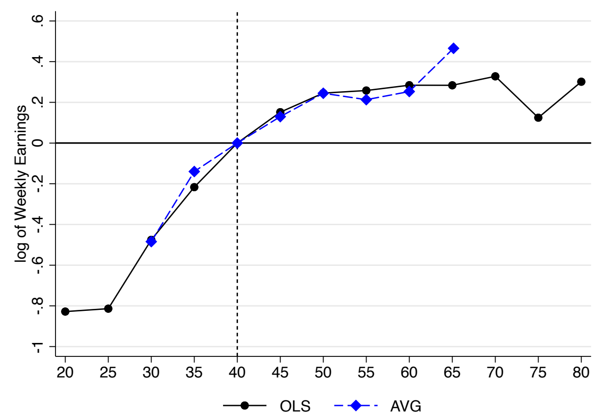


NLSY 79

(c) log Weekly Wage

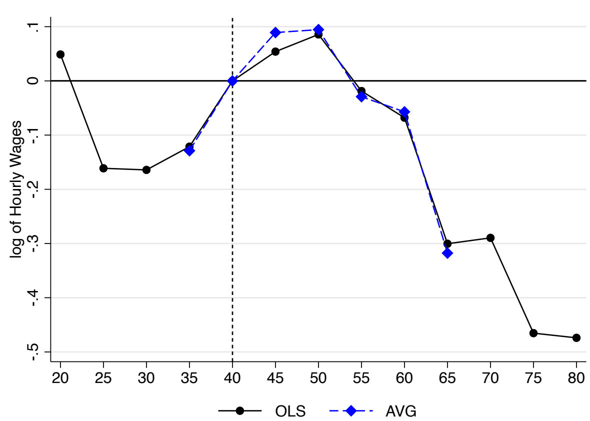


(d) log Weekly Earnings



PSID

(e) log Weekly Wage



(f) log Weekly Earnings

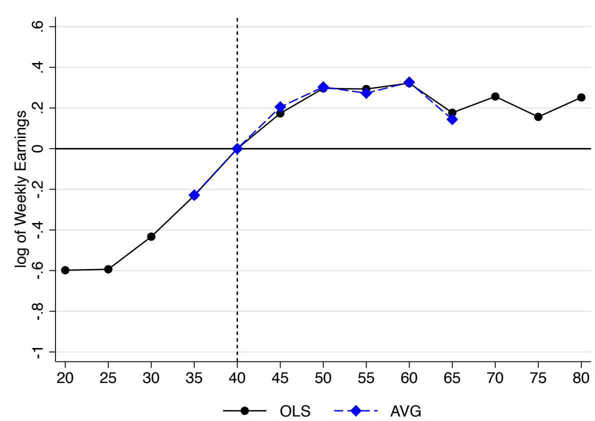
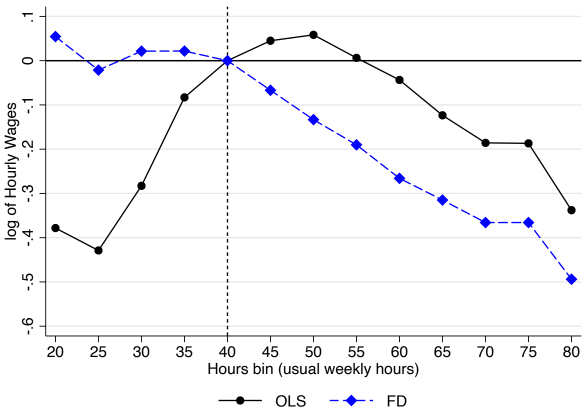


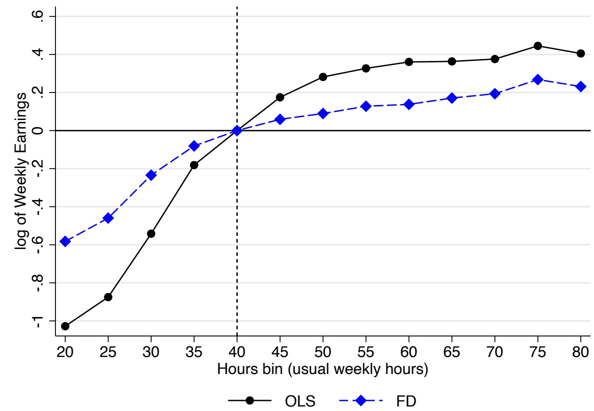
Figure A.4: Cross-Section vs. Within-Person Varation

CPS ORG

(a) log Weekly Wage

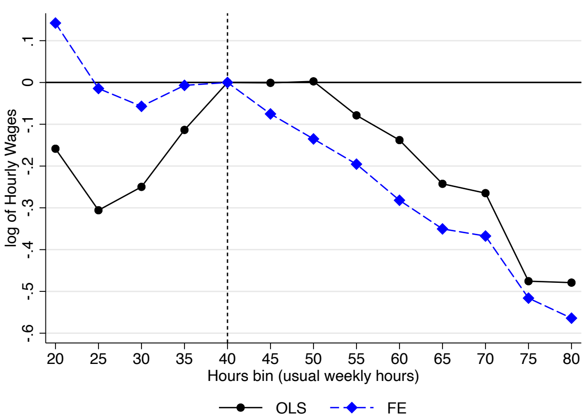


(b) log Weekly Earnings

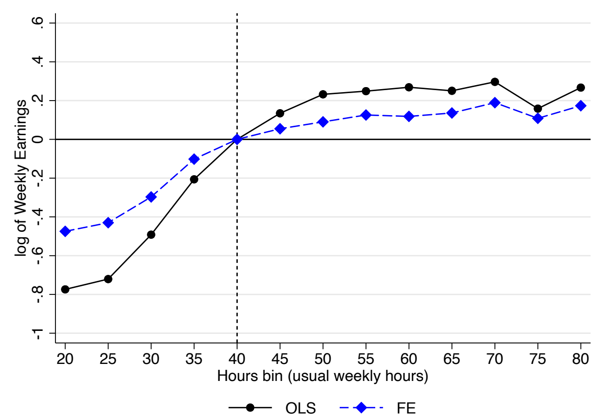


NLSY 79

(c) log Weekly Wage

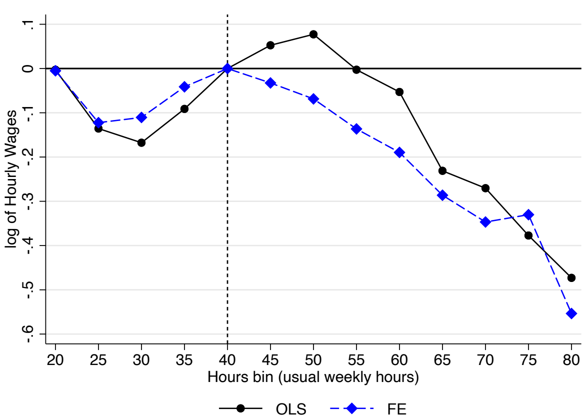


(d) log Weekly Earnings



PSID

(e) log Weekly Wage



(f) log Weekly Earnings

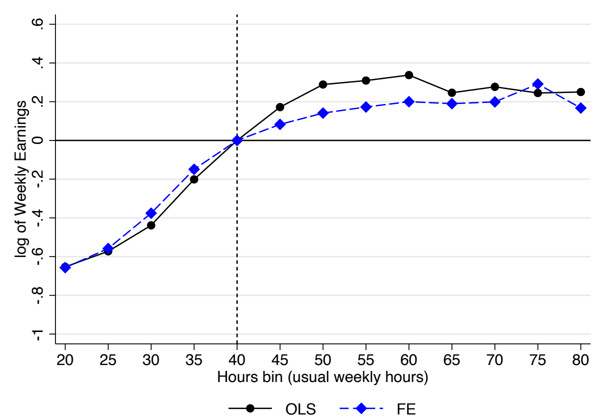
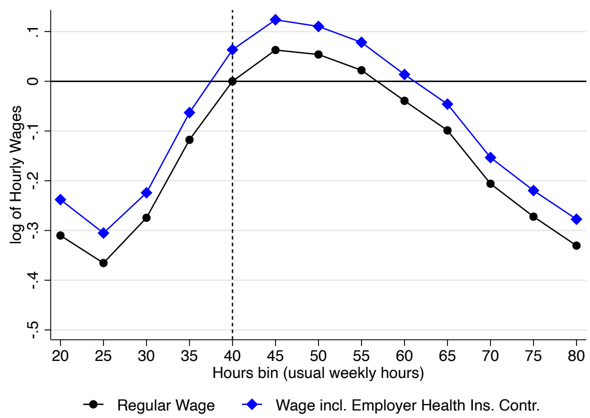
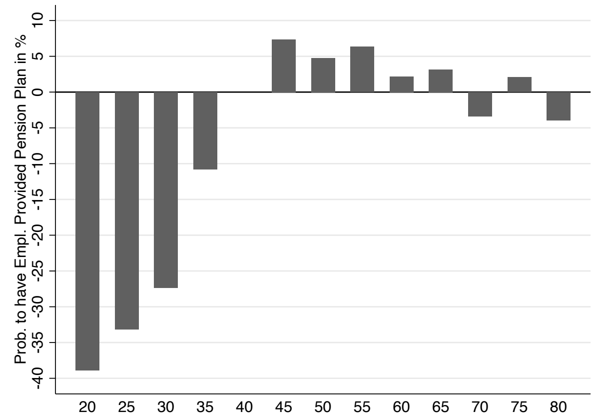


Figure A.5: The Role of Benefits (ASEC)

(a) Wages vs. Wages + Employer Contributions to Health Insurance



(b) Probability of Being Included in an Employer-Provided Pension or Retirement Plan



A.2 Tables

A.2.1 Determinants of Hours Worked in the CPS-ASEC

Table A.1: Coefficients of Different Determinants of Log Usual Weekly Hours Worked in the CPS ASEC

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
Unconditional Mean	3.709	3.746	3.774	3.804	3.716	3.749	3.777	3.804
Constant	3.692***	3.709***	3.725***	3.763***	1.855***	1.690***	1.644***	1.496***
Ages 35-44	0.002	0.010***	0.018***	0.016***	-0.009	0.004*	0.003	-0.018***
Ages 45-54	-0.003	0.005***	0.013***	0.013***	-0.004	-0.001	0.000	-0.009
Ages 55-64	-0.043***	-0.047***	-0.047***	-0.048***	-0.036***	-0.024***	-0.035***	-0.047***
Hispanic	-0.022***	-0.032***	-0.034***	-0.010	-0.022***	-0.018***	-0.007	-0.001
Black	-0.033***	-0.033***	-0.026***	-0.031***	0.002	-0.010***	-0.013**	-0.011
Married	0.038***	0.047***	0.044***	0.048***	0.041***	0.017***	0.021***	0.016**
1 Child Aged 0-4	-0.002	0.008***	0.006**	-0.001	-0.005	0.002	0.002	0.000
2 Children Aged 0-4	-0.005	0.008***	0.019***	0.028***	-0.023*	0.001	0.004	0.009
3+ Children Aged 0-4	-0.029**	-0.008	0.007	0.033*	0.005	-0.011	0.005	-0.030
1 Child Aged 5+	0.012***	0.013***	0.017***	0.010***	0.011*	0.008***	0.007*	-0.001
2 Children Aged 5+	0.012***	0.017***	0.018***	0.016***	0.002	0.012***	0.009**	0.010*
3+ Children Aged 5+	0.012***	0.017***	0.025***	0.027***	0.001	0.012***	-0.002	0.003
1st Quintile Other Inc.	0.026***	0.026***	0.022***	0.022***	0.031***	0.009***	0.009*	-0.009
2nd Quintile Other Inc.	0.001	0.016***	0.019***	-0.001	0.013**	0.003	0.013***	-0.005
4th Quintile Other Inc.	0.005	-0.006***	-0.005**	-0.023***	0.009	-0.006**	-0.004	-0.011*
5th Quintile Other Inc.	-0.002	-0.015***	-0.010***	-0.015***	0.008	-0.013***	-0.001	-0.012*
Lagged log Hours	–	–	–	–	0.489***	0.544***	0.560***	0.612***
R^2	0.019	0.027	0.028	0.023	0.289	0.304	0.336	0.407
Observations	41466	202355	70656	38430	5312	37713	14395	7859

Note: */**/*** indicate statistical significance at the 10%/5%/1% level, respectively.

Table A.2: Probability of Working Long Hours (Usual Weekly Hours $\geq$ 50) as Dependent Variable (CPS ASEC)

(a) Explantory Power of Different Determinants

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.012	0.014	0.017	0.013	0.193	0.220	0.255	0.302
Excluded Regressors								
Age	0.012	0.013	0.016	0.012	0.192	0.219	0.254	0.301
Race/Ethnicity	0.003	0.008	0.014	0.011	0.190	0.219	0.254	0.301
Marital Status	0.011	0.012	0.015	0.011	0.191	0.219	0.253	0.301
Age & # of Children	0.012	0.013	0.016	0.012	0.192	0.219	0.254	0.302
Other Income Quintile	0.012	0.013	0.016	0.012	0.192	0.220	0.254	0.302
Lagged Hours	–	–	–	–	0.014	0.012	0.021	0.014

Note: R^2 are based on a linear regression model. In the right panel, "Lagged Hours" corresponds to an indicator taking the value 1 if the individual worked at least 50 hours usually per week in the previous year.

(b) Coefficient Estimates

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
Unconditional Mean	0.142	0.213	0.305	0.397	0.141	0.208	0.301	0.390
Constant	0.171***	0.179***	0.243***	0.349***	0.042**	0.089***	0.118***	0.203***
Ages 35-44	0.006	0.008***	0.021***	0.016**	-0.001	0.014**	0.009	-0.020
Ages 45-54	-0.008	-0.004*	0.015***	0.011	-0.016	0.004	0.003	-0.014
Ages 55-64	-0.019***	-0.043***	-0.042***	-0.034***	-0.025*	-0.023***	-0.029**	-0.038**
Hispanic	-0.070***	-0.079***	-0.078***	-0.064***	-0.045***	-0.048***	-0.040**	-0.030
Black	-0.064***	-0.070***	-0.081***	-0.097***	0.006	-0.038***	-0.042***	-0.066***
Married	0.035***	0.054***	0.070***	0.069***	0.043***	0.023***	0.059***	0.044***
1 Child Aged 0-4	-0.006	0.014***	0.015***	0.012	-0.008	0.014**	0.015	0.008
2 Children Aged 0-4	-0.015	0.020***	0.051***	0.060***	-0.052*	0.014	0.026	0.026
3+ Children Aged 0-4	0.014	-0.009	0.024	0.025	0.046	-0.065*	0.008	-0.013
1 Child Aged 5+	0.009*	0.015***	0.027***	-0.006	0.008	0.007	0.022**	-0.013
2 Children Aged 5+	-0.002	0.019***	0.035***	0.028***	-0.008	0.015***	0.025**	0.015
3+ Children Aged 5+	0.009	0.024***	0.042***	0.052***	-0.007	0.024***	0.005	-0.010
1st Quintile Other Inc.	0.012**	0.018***	0.016**	0.012	0.009	0.003	-0.002	-0.007
2nd Quintile Other Inc.	0.006	0.038***	0.037***	-0.001	0.005	0.008	0.030***	-0.002
4th Quintile Other Inc.	0.009	-0.003	-0.024***	-0.050***	0.036**	-0.001	-0.011	-0.018
5th Quintile Other Inc.	0.000	0.004	-0.006	-0.005	0.007	0.000	-0.005	-0.004
Lagged Prob. $h \geq 50$	–	–	–	–	0.401***	0.445***	0.483***	0.536***
R^2	0.012	0.014	0.017	0.013	0.193	0.220	0.255	0.302
Observations	41466	202355	70656	38430	5312	37713	14395	7859

Note: */**/*** indicate statistical significance at the 10%/5%/1% level, respectively.

Table A.3: Explantory Power of Different Determinants of Hours Worked in the CPS ASEC incl. Dummies for 3-digit Occupation

(a) Dependent Variable — Log Usual Weekly Hours

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.102	0.112	0.129	0.150	0.348	0.339	0.384	0.451
Excluded Regressors								
Age	0.097	0.107	0.123	0.144	0.343	0.337	0.381	0.448
Race/Ethnicity	0.101	0.111	0.128	0.150	0.347	0.339	0.384	0.451
Marital Status	0.098	0.109	0.126	0.148	0.342	0.339	0.383	0.451
Age & # of Children	0.101	0.112	0.128	0.150	0.347	0.339	0.384	0.451
Other Income Quintile	0.099	0.109	0.126	0.148	0.345	0.338	0.383	0.451
Occupation	0.019	0.027	0.028	0.023	0.289	0.304	0.336	0.407
Lagged Hours	–	–	–	–	0.159	0.122	0.148	0.160

(b) Dependent Variable — Probability of Working Long Hours (Usual Weekly Hours ≥ 50)

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.091	0.090	0.093	0.123	0.256	0.256	0.295	0.348
Excluded Regressors								
Age	0.091	0.089	0.092	0.121	0.256	0.255	0.294	0.348
Race/Ethnicity	0.087	0.087	0.092	0.121	0.255	0.255	0.295	0.348
Marital Status	0.090	0.089	0.092	0.122	0.255	0.256	0.294	0.348
Age & # of Children	0.091	0.090	0.093	0.122	0.256	0.256	0.295	0.348
Other Income Quintile	0.091	0.089	0.092	0.122	0.256	0.256	0.294	0.348
Occupation	0.012	0.014	0.017	0.013	0.193	0.220	0.255	0.302
Lagged Hours	–	–	–	–	0.126	0.097	0.114	0.133

A.2.2 Determinants of Hours Worked in the CPS-ASEC When Using the Same Sample in the Regressions With and Without Lagged Hours

Table A.4: Explantory Power of Different Determinants of Log Usual Weekly Hours Worked in the CPS ASEC Using the Same Sample in Both Table Panels

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.030	0.024	0.030	0.027	0.289	0.304	0.336	0.407
Excluded Regressors								
Age	0.023	0.019	0.020	0.015	0.285	0.302	0.333	0.402
Race/Ethnicity	0.026	0.021	0.029	0.026	0.287	0.303	0.336	0.407
Marital Status	0.017	0.019	0.023	0.024	0.283	0.303	0.335	0.406
Age & # of Children	0.029	0.022	0.028	0.026	0.288	0.303	0.336	0.406
Other Income Quintile	0.025	0.020	0.026	0.025	0.286	0.303	0.335	0.406
Lagged Hours	–	–	–	–	0.030	0.024	0.030	0.027

Notes: The right table panel is identical to the right table panel of Table 1. The left panel table reports results for using the same sample as used in the right panel, which is a subsample of the one used in the left panel of Table 1.

Table A.5: Explantory Power of Different Determinants of Probability of Working Long Hours (Usual Weekly Hours ≥ 50) as Dependent Variable in the CPS ASEC Using the Same Sample in Both Table Panels

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.014	0.012	0.021	0.014	0.193	0.220	0.255	0.302
Excluded Regressors								
Age	0.013	0.011	0.020	0.013	0.192	0.219	0.254	0.301
Race/Ethnicity	0.007	0.008	0.020	0.012	0.190	0.219	0.254	0.301
Marital Status	0.011	0.010	0.016	0.011	0.191	0.219	0.253	0.301
Age & # of Children	0.014	0.011	0.020	0.013	0.192	0.219	0.254	0.302
Other Income Quintile	0.013	0.011	0.018	0.013	0.192	0.220	0.254	0.302
Lagged Hours	–	–	–	–	0.014	0.012	0.021	0.014

Notes: The right table panel is identical to the right table panel of Table A.2a. The left panel table reports results for using the same sample as used in the right panel, which is a subsample of the one used in the left panel of Table A.2a.

Table A.6: Explantory Power of Different Determinants of Log Usual Weekly Hours Worked in the CPS ASEC Incl. Dummies for 3-digit Occupation Using the Same Sample in Both Table Panels

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.159	0.122	0.148	0.160	0.348	0.339	0.384	0.451
Excluded Regressors								
Age	0.151	0.118	0.141	0.151	0.343	0.337	0.381	0.448
Race/Ethnicity	0.158	0.121	0.147	0.159	0.347	0.339	0.384	0.451
Marital Status	0.149	0.120	0.144	0.159	0.342	0.339	0.383	0.451
Age & # of Children	0.158	0.120	0.147	0.159	0.347	0.339	0.384	0.451
Other Income Quintile	0.155	0.119	0.145	0.159	0.345	0.338	0.383	0.451
Occupation	0.030	0.024	0.030	0.027	0.289	0.304	0.336	0.407
Lagged Hours	–	–	–	–	0.159	0.122	0.148	0.160

Notes: The right table panel is identical to the right table panel of Table A.3a. The left panel table reports results for using the same sample as used in the right panel, which is a subsample of the one used in the left panel of Table A.3a.

Table A.7: Explantory Power of Different Determinants of Probability of Working Long Hours (Usual Weekly Hours ≥ 50) as Dependent Variable in the CPS ASEC Incl. Dummies for 3-digit Occupation Using the Same Sample in Both Table Panels

	No Control for Lagged Hours				Control for Lagged Hours			
	LHS	HS	Bach	Bach+	LHS	HS	Bach	Bach+
All Regressors	0.126	0.097	0.114	0.133	0.256	0.256	0.295	0.348
Excluded Regressors								
Age	0.125	0.096	0.113	0.132	0.256	0.255	0.294	0.348
Race/Ethnicity	0.123	0.095	0.113	0.132	0.255	0.255	0.295	0.348
Marital Status	0.123	0.096	0.111	0.132	0.255	0.256	0.294	0.348
Age & # of Children	0.125	0.096	0.113	0.133	0.256	0.256	0.295	0.348
Other Income Quintile	0.125	0.096	0.111	0.133	0.256	0.256	0.294	0.348
Occupation	0.014	0.012	0.021	0.014	0.193	0.220	0.255	0.302
Lagged Hours	–	–	–	–	0.126	0.097	0.114	0.133

Notes: The right table panel is identical to the right table panel of Table A.3b. The left panel table reports results for using the same sample as used in the right panel, which is a subsample of the one used in the left panel of Table A.3b.

A.2.3 Determinants of Hours Worked in the SCF

Table A.8: Coefficients of Different Determinants of Log Usual Weekly Hours Worked in the SCF

	LHS	HS	Bach	Bach+
Unconditional Mean	3.753	3.775	3.778	3.817
Constant	3.795***	3.767***	3.756***	3.773***
Ages 35-44	0.008	-0.004	0.007	-0.009**
Ages 45-54	-0.004	-0.013**	0.009*	-0.011**
Ages 55-64	-0.030***	-0.032***	-0.035***	-0.020***
Hispanic	-0.028***	-0.045***	-0.048***	-0.027***
Black	-0.047***	-0.053***	-0.035***	-0.005
Married	-0.011	0.021***	0.047***	0.047***
1 Child Aged 0-4	-0.005	-0.001	0.005	-0.003
2 Children Aged 0-4	-0.026*	-0.022**	0.009	0.011
3+ Children Aged 0-4	-0.024	0.001	0.082**	0.030
1 Child Aged 5+	0.003	0.028***	0.007	0.007
2 Children Aged 5+	0.031***	0.016***	0.024***	0.019***
3+ Children Aged 5+	0.017*	0.025***	0.027***	0.042***
1st Quintile Other Inc.	0.035***	0.039***	0.052***	0.041***
2nd Quintile Other Inc.	-0.006	-0.015***	0.025***	0.023***
4th Quintile Other Inc.	-0.014	-0.015***	0.005	-0.008*
5th Quintile Other Inc.	-0.004	-0.008	0.016**	0.015***
log Wealth-Income Ratio	0.007***	-0.005***	-0.003*	0.008***
R^2	0.041	0.032	0.034	0.024
Observations	5382	15016	12410	21407

Note: */**/*** indicate statistical significance at the 10%/5%/1% level, respectively.

Table A.9: Probability of Working Long Hours (Usual Weekly Hours ≥ 50) as Dependent Variable (SCF)

(a) Explantory Power of Different Determinants

	LHS	HS	Bach	Bach+
All Regressors	0.036	0.024	0.036	0.016
Excluded Regressors				
Age	0.033	0.023	0.033	0.016
Race/Ethnicity	0.029	0.016	0.030	0.015
Marital Status	0.036	0.024	0.032	0.015
Age & # of Children	0.033	0.022	0.031	0.014
Other Income Quintile	0.034	0.019	0.025	0.014
log Wealth-Income Ratio	0.034	0.024	0.036	0.013

Note: R^2 are based on a linear regression model. In the right panel, "Lagged Hours" corresponds to an indicator taking the value 1 if the individual worked at least 50 hours usually per week in the previous year.

(b) Coefficient Estimates

	LHS	HS	Bach	Bach+
Unconditional Mean	0.202	0.268	0.280	0.403
Constant	0.250***	0.279***	0.248***	0.340***
Ages 35-44	-0.003	-0.025**	-0.037***	-0.012
Ages 45-54	-0.027	-0.034***	-0.012	0.000
Ages 55-64	-0.082***	-0.051***	-0.084***	-0.009
Hispanic	-0.070***	-0.097***	-0.096***	-0.080***
Black	-0.089***	-0.106***	-0.080***	-0.041***
Married	0.008	0.024**	0.091***	0.059***
1 Child Aged 0-4	-0.010	0.018	-0.008	0.014
2 Children Aged 0-4	0.014	-0.041*	-0.002	0.062***
3+ Children Aged 0-4	-0.258***	0.067	0.328***	0.129*
1 Child Aged 5+	0.007	0.031***	0.032***	0.028***
2 Children Aged 5+	0.027	0.039***	0.071***	0.036***
3+ Children Aged 5+	0.005	0.048***	0.087***	0.071***
1st Quintile Other Inc.	0.025	0.053***	0.130***	0.040***
2nd Quintile Other Inc.	-0.007	-0.037***	0.051***	0.019
4th Quintile Other Inc.	-0.049**	-0.024**	-0.003	-0.018
5th Quintile Other Inc.	-0.038	-0.032**	0.061***	0.032***
log Wealth-Income Ratio	0.013***	-0.003	-0.001	0.024***
R^2	0.036	0.024	0.036	0.016
Observations	5382	15016	12410	21407

Note: */**/*** indicate statistical significance at the 10%/5%/1% level, respectively.

A.2.4 Determinants of Hours Worked in the SCF by Age

Dependent Variable Log Usual Weekly Hours Worked

Table A.10: Ages 25-34

	LHS	HS	Bach	Bach+
All Regressors	0.132	0.058	0.066	0.065
Excluded Regressors				
Race/Ethnicity	0.126	0.045	0.056	0.063
Marital Status	0.127	0.056	0.063	0.063
Age & # of Children	0.124	0.051	0.056	0.060
Other Income Quintile	0.092	0.046	0.051	0.038
log Wealth-Income Ratio	0.128	0.056	0.065	0.064

Table A.11: Ages 35-44

	LHS	HS	Bach	Bach+
All Regressors	0.105	0.029	0.040	0.049
Excluded Regressors				
Race/Ethnicity	0.101	0.021	0.040	0.046
Marital Status	0.105	0.028	0.035	0.047
Age & # of Children	0.089	0.024	0.033	0.037
Other Income Quintile	0.083	0.020	0.031	0.044
log Wealth-Income Ratio	0.093	0.028	0.040	0.040

Table A.12: Ages 45-54

	LHS	HS	Bach	Bach+
All Regressors	0.066	0.027	0.051	0.047
Excluded Regressors				
Race/Ethnicity	0.060	0.016	0.035	0.045
Marital Status	0.065	0.026	0.048	0.044
Age & # of Children	0.055	0.024	0.045	0.039
Other Income Quintile	0.047	0.017	0.045	0.040
log Wealth-Income Ratio	0.065	0.027	0.051	0.045

Table A.13: Ages 55-64

	LHS	HS	Bach	Bach+
All Regressors	0.134	0.058	0.094	0.081
Excluded Regressors				
Race/Ethnicity	0.096	0.052	0.079	0.079
Marital Status	0.134	0.058	0.081	0.050
Age & # of Children	0.132	0.029	0.092	0.077
Other Income Quintile	0.111	0.046	0.079	0.065
log Wealth-Income Ratio	0.134	0.057	0.068	0.073

Dependent Variable: Probability of Working Long Hours

Table A.14: Ages 25-34

	LHS	HS	Bach	Bach+
All Regressors	0.150	0.041	0.075	0.060
Excluded Regressors				
Race/Ethnicity	0.130	0.035	0.065	0.059
Marital Status	0.144	0.040	0.072	0.059
Age & # of Children	0.140	0.034	0.065	0.048
Other Income Quintile	0.118	0.036	0.047	0.050
log Wealth-Income Ratio	0.149	0.041	0.074	0.059

Table A.15: Ages 35-44

	LHS	HS	Bach	Bach+
All Regressors	0.083	0.036	0.046	0.041
Excluded Regressors				
Race/Ethnicity	0.080	0.028	0.045	0.039
Marital Status	0.082	0.035	0.043	0.041
Age & # of Children	0.075	0.030	0.032	0.033
Other Income Quintile	0.070	0.026	0.034	0.038
log Wealth-Income Ratio	0.072	0.036	0.046	0.033

Table A.16: Ages 45-54

	LHS	HS	Bach	Bach+
All Regressors	0.082	0.024	0.045	0.029
Excluded Regressors				
Race/Ethnicity	0.074	0.015	0.034	0.027
Marital Status	0.082	0.024	0.043	0.025
Age & # of Children	0.046	0.019	0.041	0.025
Other Income Quintile	0.076	0.017	0.039	0.026
log Wealth-Income Ratio	0.080	0.024	0.045	0.026

Table A.17: Ages 55-64

	LHS	HS	Bach	Bach+
All Regressors	0.145	0.088	0.092	0.060
Excluded Regressors				
Race/Ethnicity	0.131	0.052	0.071	0.058
Marital Status	0.143	0.085	0.080	0.037
Age & # of Children	0.131	0.052	0.086	0.059
Other Income Quintile	0.085	0.084	0.059	0.052
log Wealth-Income Ratio	0.142	0.084	0.075	0.060

A.2.5 Determinants of Hours Worked in the SCF by Age and Using log Wealth instead of the log Wealth to Income Ratio

Dependent Variable Log Usual Weekly Hours Worked

Table A.18: Ages 25-34

	LHS	HS	Bach	Bach+
All Regressors	0.144	0.059	0.082	0.072
Excluded Regressors				
Race/Ethnicity	0.139	0.049	0.075	0.070
Marital Status	0.139	0.058	0.079	0.070
Age & # of Children	0.136	0.051	0.072	0.068
Other Income Quintile	0.107	0.046	0.062	0.044
log Wealth	0.128	0.057	0.065	0.064

Table A.19: Ages 35-44

	LHS	HS	Bach	Bach+
All Regressors	0.114	0.035	0.050	0.077
Excluded Regressors				
Race/Ethnicity	0.108	0.029	0.050	0.075
Marital Status	0.113	0.034	0.047	0.076
Age & # of Children	0.097	0.029	0.044	0.068
Other Income Quintile	0.077	0.023	0.039	0.069
log Wealth	0.102	0.029	0.039	0.040

Table A.20: Ages 45-54

	LHS	HS	Bach	Bach+
All Regressors	0.077	0.033	0.060	0.081
Excluded Regressors				
Race/Ethnicity	0.072	0.026	0.047	0.080
Marital Status	0.076	0.033	0.059	0.079
Age & # of Children	0.066	0.031	0.055	0.073
Other Income Quintile	0.053	0.020	0.052	0.070
log Wealth	0.065	0.027	0.051	0.045

Table A.21: Ages 55-64

	LHS	HS	Bach	Bach+
All Regressors	0.146	0.063	0.070	0.082
Excluded Regressors				
Race/Ethnicity	0.117	0.060	0.057	0.079
Marital Status	0.146	0.063	0.059	0.055
Age & # of Children	0.143	0.031	0.067	0.078
Other Income Quintile	0.122	0.046	0.059	0.061
log Wealth	0.134	0.057	0.068	0.073

Dependent Variable Probability of Working Long Hours

Table A.22: Ages 25-34

	LHS	HS	Bach	Bach+
All Regressors	0.155	0.046	0.084	0.066
Excluded Regressors				
Race/Ethnicity	0.139	0.042	0.076	0.066
Marital Status	0.150	0.045	0.081	0.065
Age & # of Children	0.145	0.039	0.074	0.055
Other Income Quintile	0.125	0.040	0.054	0.056
log Wealth	0.149	0.042	0.075	0.060

Table A.23: Ages 35-44

	LHS	HS	Bach	Bach+
All Regressors	0.094	0.041	0.053	0.065
Excluded Regressors				
Race/Ethnicity	0.091	0.035	0.053	0.064
Marital Status	0.093	0.041	0.051	0.064
Age & # of Children	0.087	0.036	0.040	0.059
Other Income Quintile	0.076	0.031	0.039	0.063
log Wealth	0.072	0.036	0.045	0.033

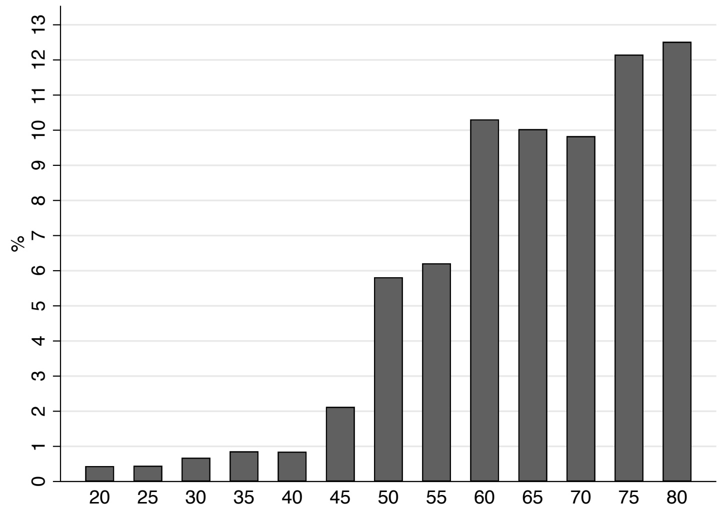
Table A.24: Ages 45-54

	LHS	HS	Bach	Bach+
All Regressors	0.093	0.026	0.048	0.066
Excluded Regressors				
Race/Ethnicity	0.087	0.019	0.040	0.065
Marital Status	0.093	0.026	0.047	0.063
Age & # of Children	0.058	0.021	0.045	0.061
Other Income Quintile	0.084	0.018	0.043	0.059
log Wealth	0.080	0.024	0.044	0.026

Table A.25: Ages 55-64

	LHS	HS	Bach	Bach+
All Regressors	0.162	0.084	0.077	0.079
Excluded Regressors				
Race/Ethnicity	0.154	0.052	0.059	0.077
Marital Status	0.160	0.082	0.066	0.060
Age & # of Children	0.145	0.047	0.072	0.078
Other Income Quintile	0.094	0.080	0.047	0.065
log Wealth	0.142	0.084	0.075	0.060

Figure B.1: Probability of being Top-Coded by Usual Weekly Hours bin for Men



Sample Period: Sep 1995-August 2007

B The Role of Top-Coding

Figure 3b of the main text showed that, for our sample of men in the ORG, mean earnings were relatively flat in usual weekly hours beyond 50 hours per week. In this section we attempt to analyze the quantitative role of top-coding for this pattern. To see why top-coding could potentially be relevant, consider an extreme case where no one working less than 50 hours is top coded, everyone from 50 hours on is top-coded, and all top-coded earnings are replaced with a single value. In this case, even if true earnings were increasing in hours, observed earnings would be completely flat beyond 50. The following paragraphs provide suggestive evidence against this possiblity, i.e. we conclude that top-coding is not the major driver of the relatively flat earnings-hours relationship beyond 50 hours. One partial exception is that those with graduate degrees have a higher incidence of top-coding, and for these workers top-coding may modestly dampen the rate at which earnings increase in the long hours region. But this is only the case in the CPS ORG because of the specific top-coding procedure, and not in the other datasets.

The sample for our analysis starts in September 1995, the first months from which onwards IPUMS provide information whether earnings have been imputed or not in the ORG. Between September 1995 and December 1996 earnings were top-coded at $1,923 per week (corresponding to $100,000 per year assuming 52 weeks of work) in nominal terms. Since January 1998, earnings have been toop-coded at $2,885.61 per week (corresponding to $150,000 per year assuming 52 weeks of work) in nominal terms. Figure B.1 shows results for our sample of men age 25-64. Below 45 hours, top-coding is negligible and even in the 45-49 hours bin the earnings of only 2% of men are subject to top-coding. From 50 hours onwards, the probability of earnings being top-coded becomes more prevalent and increases in usual hours worked, although not monotonically.

Our first step is to compare results in the CPS ORG and ASEC, using the same years and sample criteria for the ASEC as in our ORG sample.[28] In contrast to the ORG, the nominal top-codes in the ASEC are regularly adjusted and are generally higher. As one might expect, this leads to a lower probability of being top-coded in ASEC than in ORG, as seen in Figure B.2a. In addition to different top-code thresholds,

[28]To be precise, the sample period for the ASEC is 1996 through 2008 since hours and earnings are reported for the previous year.

Figure B.2: Different Top Codes in CPS ORG and ASEC for Men (1995-2007)

(a) Top-Coding Probability (b) Wages-Hours Relationship

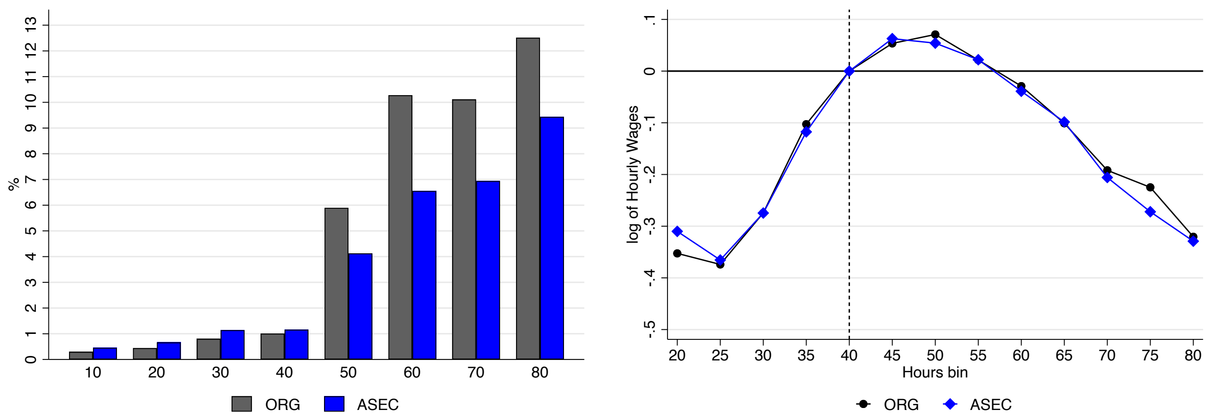
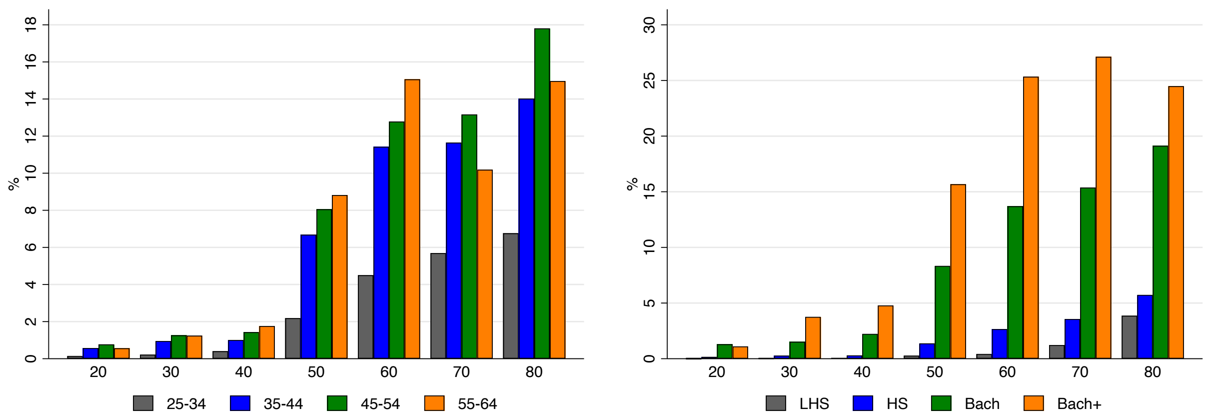


Figure B.3: Top-Coding Probabilities in ORG for Men (1995-2007)

(a) By Age (b) By Education



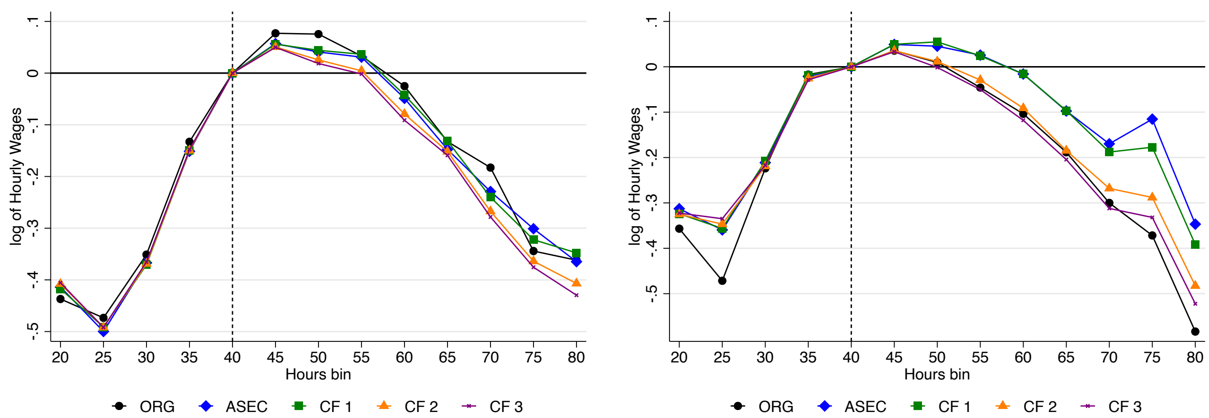
ORG and ASEC also differ in how earnings are assigned to top-coded individuals. In the ORG, top-coded individuals are assigned the top-code. In contrast, until 2011 in the ASEC top-coded individuals were assigned the mean earnings of the top-coded. Specifically, the means earnings were calculated and assigned by cells defined by gender, race (black vs. hispanic vs. rest) and labor supply (full-year-full-time workers, i.e. weeks worked $\geq$ 50 and weekly hours $\geq$ 35, vs. rest). Figure B.2b shows that despite the different top-coding procedures, the aggregate wage-hours relationship is virtually identical. This is consistent with the notion that top-coding is not a major issue in the aggregate.

Next, we analyze the role of top-coding among specific groups of workers. Figure B.3 shows the probability of being top-coded in ORG by age and education. The probability of top-coding is increasing in age up to the 60 hours bin, although the differences are relatively small beyond age 34. The probability of top-coding is strongly increasing in education, and peaks around 25% of workers with a graduate degree working at least 60 hours.

Figure B.4 plots the cross-sectional wage-hours relationship for the ORG and ASEC for the age and education group with the overall highest probability of being top-coded: men aged 55-64, and men with a

Figure B.4: Comparing Top-Coding Procedures: ORG vs. ASEC

(a) Age Group 55-64 (b) Bach+



graduate degree. In addition, we also analyze the following counterfactual top-coding procedures using the ASEC:

CF 1. Impose ORG top-code threshold, replace top-coded with average earnings of top-coded by race and labor supply.

CF 2. Keep ASEC top-code threshold, replace top-coded with ASEC top-code

CF 3. Impose ORG top-code threshold, replace top-coded with ORG top-code

Counterfactual 1 is informative about how important a more binding top-code is, holding fixed the top-coding replacement strategy in ASEC. Counterfactual 2 is informative about how important the replacement strategy of top-coded values is, holding fixed the top-code in ASEC. Countefactual 3 is informative about the combination of Counterfactuals 1-2 together.

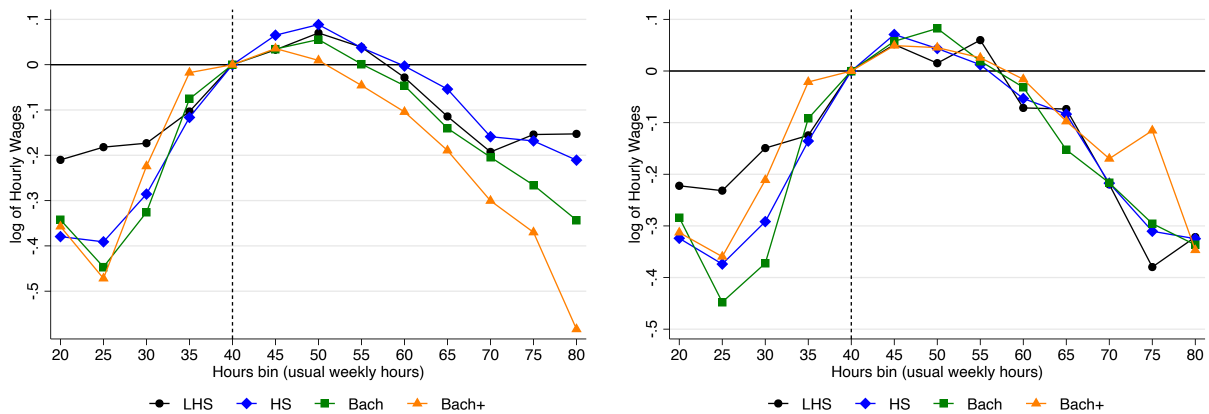
Figure B.4a shows that for the age group 55-64 all wage-hours profiles look very similar. This suggests, similar to the aggregate pattern, that the more restricted top-coding in the ORG does not have important effects when distinguishing between age groups.

By contrast, in Figure B.4b, we observe noticeably different wage-hours profiles for men with a graduate degree in the ORG vs. the ASEC. Specifically, in the 50 hour bin the average hourly wage in the ORG is 5 log points below the ASEC; in the 60 hour bin this difference has increased to 9 log points. When we replace both the ASEC top-code threshold with the ORG threshold, and the ASEC top-coding replacement procedure with the ORG replacement procedure (this can be seen by comparing CF 3 and the ORG profile), we find nearly identical results to the actual ORG results. The main reason for the difference between the ASEC profile and the ORG profile is thus not the lower top-coding threshold (this can be seen by comparing CF 1 with the ASEC profile). Instead, the major source of the difference is the difference in the replacement strategies (this can be seen by comparing CF 2 and the ASEC profile).

Figure B.5a shows again the patterns by education in the ORG from the main text, from which one can see that the profile is more depressed for those with a Bachelor and a graduate degree (Bach+). Figure B.4b suggests that some of this pattern might be related to top-coding. Figure B.5b shows the patterns by education for the ASEC, where the profiles lie mostly on top of each other. Hence, while the gaps by education in the ORG may partly reflect top-coding, the ASEC results are in line with our main interpretation on the role of top-coding, namely that top-coding is not the main driver of our finding.

Figure B.5: Cross-Sectional Relationship between Wages and Hours: ORG vs. ASEC

(a) ORG (b) ASEC



We conclude this section by addressing a final potential issue, which is that if true earnings above the top-code are increasing in hours worked, then replacing the top-coded earnings of all long hours workers with the same value could flatten the earnings profile among these workers. (Recall that the replacement values for the top-coded in the ASEC did account for whether workers worked at least 35 hours per week, but did not distinguish between, for example, workers who worked 50 hours per week and those who worked 60 hours per week). To address this, we turn to the PSID. Since the mid-nineties, the PSID's top-code for wage earnings of the household head is \$10 million. In fact, this threshold is so high that no one in the PSID satisfying our sample selection criteria is top-coded. Given the small sample size in the PSID, the following exercise will be for the years 1996-2018. Similar to the previous counterfactuals, we know impose the ASEC top-coding strategy on the PSID and compare this to the actual PSID withou no top-coding.[29] Figure B.6 shows results for the aggregate as well as for those with a college degree (for sample size reasons we do not distinguish between a bachelor and graduate degree). While the PSID shows slightly different patterns than ASEC, the main take-away is that imposing the ASEC top coding strategy yields very similar results to the actual PSID which effectively had no top-coding. This is consistent with the notion that earnings among top-coded workers do not vary strongly with hours worked above 50.

[29]When implementing this strategy, we focus on the top-codes for inclongj in ASEC which is the dominant income measure for wage and salary earners. For sample size reasons, we also group top-coded individuals only by whether someone is a full-year-full-time worker but not on race.

Figure B.6: Comparing Top-Coding Procedures: PSID vs. ASEC (1995-2018)

(a) Aggregate (b) College Degree

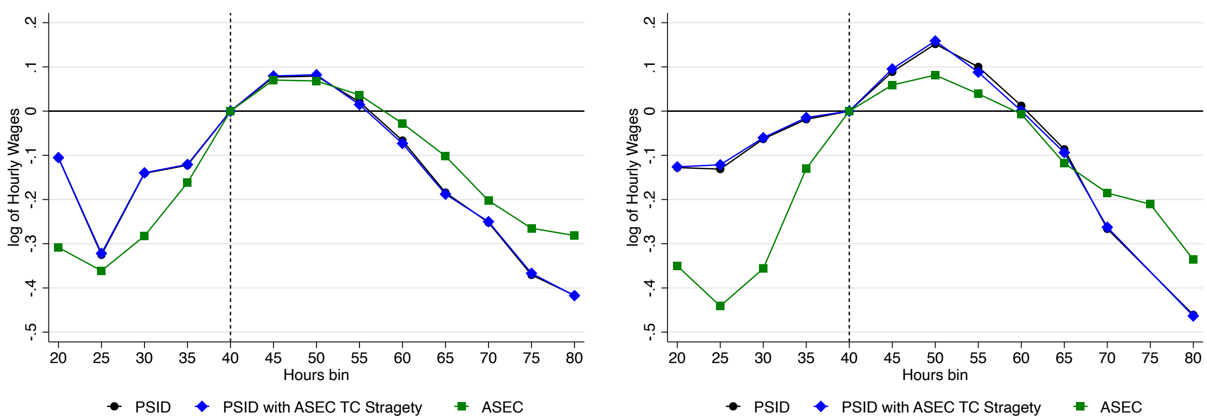
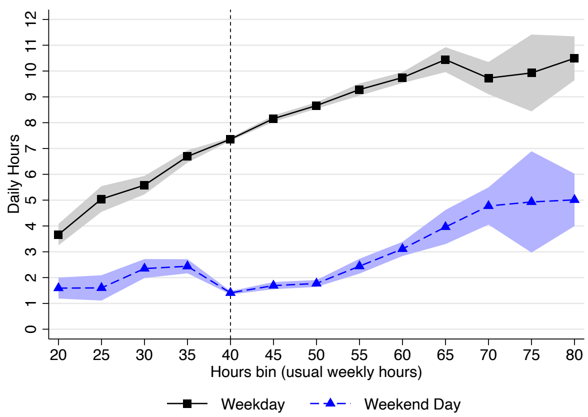


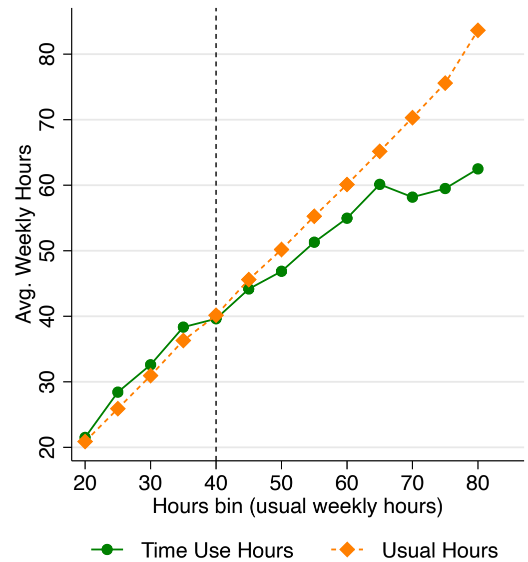
Figure C.1: Average Time Use Hours

(a) Avg. Daily Time Use Hours



Note: Shaded areas are 95% Confidence Intervals.

(b) Avg. Weekly Time Use Hours



C The Role of Measurement Error in Hours

In Figure 3b of the main text, mean earnings were relatively flat in usual weekly hours beyond 50 hours per week. If people with high hours tend to be people who have over-reported their hours, then this will artificially lead to a flatter pattern even if true earnings are increasing in hours. In this section we attempt to analyze the quantitative role of measurement error in hours for this pattern.

To assess the impact of measurement error we link observations between the CPS ORG and the American Time Use Survey (ATUS). Since 2003 the ATUS collects a time diary for a sample of individuals (not households) 2 to 5 months after their 8th CPS interview. The diary records all activities between 4am of the day preceding the ATUS interview and 4am of the interview day. It records the type of activity, starting and end point as well the location it took place. IPUMS provides a variable that aggregates these activities into "hours spent working on the main job". Importantly, this variable does not include commuting or social activities around work like a lunch break or dinner. From the last CPS interview, we also know usual hours worked, which maybe updated by the respondent at the time of the ATUS interview.

For our analysis we use the same sample restrictions as laid out in Section 2.1, but impose two further restrictions. First, the ATUS provides a variable about the interviewer's perception of data quality indicating whether or not interviewers believe the data from a particular interview should be used. Reasons for why an interview should not be used are if the interviewer thinks that the respondent intentionally provided a wrong answer, could not correctly remember activities, deliberately reported very long durations, or some other reason. We only use interviews which the interviewers suggest to use. Second, because we are interested in usual hours worked, we drop all individuals who did not work at all in the last 7 days. For example, consider someone who was an entire week on vacation and therefore reports zero hours in the time use diary. This zero is simply not informative about the person's usual hours worked, or more precisely for the usual hours worked of people with similar characteristics. Finally, to ensure a sufficiently large sample size we use all years for which the ATUS is available, i.e. 2003 through 2018.

Given our sample, our analysis proceeds as follows. We group individuals by their usual hours bin as

Table C.1: Differences between Avg. Weekly Time Use and Usual Hours by Usual Hours Bin

Usual Hours Bin	Hours Difference Levels	Percent
20-24	0.6	3.0
25-29	2.5	9.7
30-34	1.7	5.4
35-39	2.0	5.6
40-44	-0.5	-1.3
45-49	-1.4	-3.1
50-54	-3.3	-6.6
55-59	-4.0	-7.2
60-64	-5.1	-8.6
65-69	-5.0	-7.7

reported in the CPS ORG. Next, we calculate the average ATUS hours worked on a weekday and on a weekend day, respectively, for each ORG hours bin. We report these reults in Figure C.1a. Average daily time use hours on a week day increase monotonically up to the up the 65-69 usual hours bin and flatten out subsequently. Individuals reporting usual hours in the 40-44 hours bin report slightly more than 7 hours of work on a week day based on the time use data. Individuals reporting usual hours in the 65-69 hours bin report more than 10 hours of work on a week day based on the time use data. Average daily time use hours on a weekend day are slightly above 1 hour for workers whose usual weekly hours are less than 50, and increase to close to 4 hours in the 65-69 usual hours bin. Taking the time use hours at face value, this provides clear evidence that actual hours worked are increasing in reported usual hours worked.

We next compute a synthetic measure of weekly hours worked using the ATUS data, and compare it to the reported usual weekly hours in the ORG. To do so, for each usual weekly hours bin we multiply the average daily time use hours on a weekday by 5 and on a weekend day by 2, then sum the two numbers. Figure C.1b displays the results. In what follows we will focus our attention on individuals with reported usual hours below 70. As we noted earlier, there are very few individuals with reported usual hours of 70 or more, and our estimation exercise does not use any information for these individuals. Moreover, Figure C.1b indicates that the discrepancy between the two measures becomes very large for these workers, thereby casting some doubt on the reliability of these responses. On average, workers who report usual hours in the CPS in in the 30-40 hours region tend to report slightly higher actual hours in the ATUS, while workers who report usual hours in the CPS in the 45-69 hours region tend to report slightly lower actual hours in the ATUS. Table C.1 reports the magnitude of the level and percentage differences across the hours distribution. For usual hours in the 50-69 hours region the gap varies in a relatively tight range of roughly 3-5 hours, corresponding to a percentage gap between roughly 6.6% and 8.6%.

Taking the ATUS measure at face value, this evidence suggests that workers in the long hours region tend to overreport their hours. We are particularly interested in assessing the possibility that overreporting of hours in the long hours region has a large effect on our estimate of the cross-sectional wage-hours profile. To pursue this we implement the following exercise. Because the gap in the two measures is relatively constant in the 3-5 hours range over the 50-69 hours range, and most people report usual hours ending in either a 0 or a 5, we assume that all individuals with usual hours in this region have true hours in the immediate lower bin, i.e., an individual with reported usual hours in the 50-54 hours bin is now placed in the 45-49 hours bin etc... We then repeat our benchmark regression exercise to uncover the earnings-hours and wage-hours profile.

Figure C.2: Cross-Sectional Relationship between Earnings/Wages and Reported and Adjusted Hours

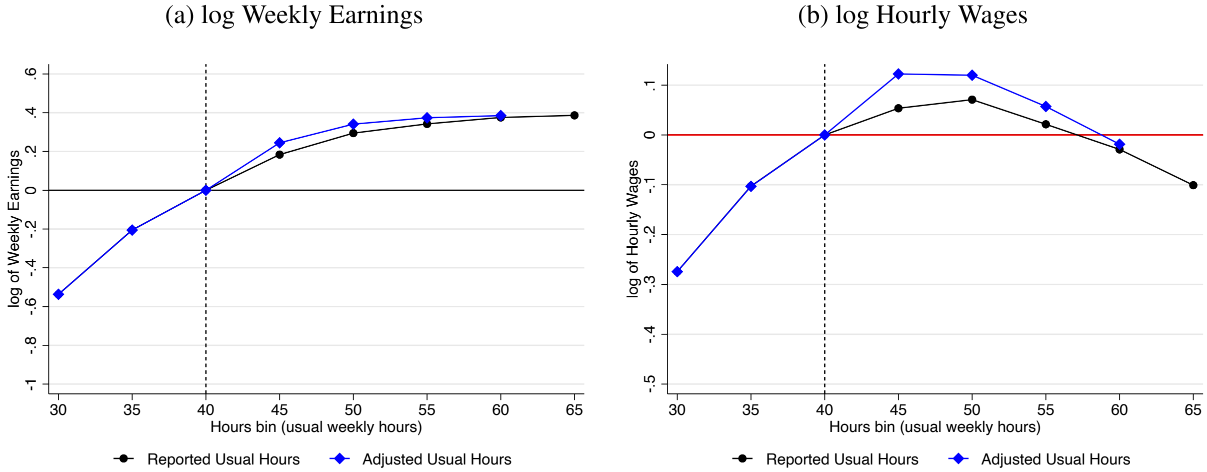


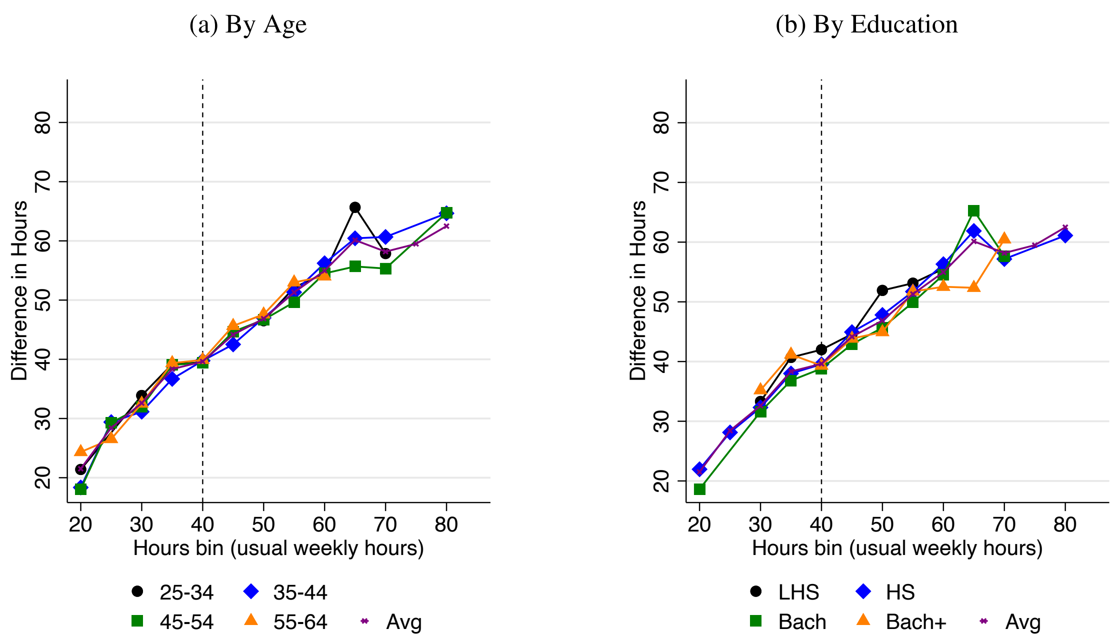
Figure C.2a shows how this affects the estimated earnings-hours profile. For comparison we have included our baseline estimate from Figure 3 in the main text. The key message is that this has virtually no impact on the finding that earnings are effectively flat in the 50-69 hours region. This should not be surprising—given that earnings are effectively flat in the CPS and our estimate of overreporting is relatively constant, the slope will be relatively unaffected.

To estimate the wage-hours profile it is not sufficient to just identify which 5 hour bin an individual belongs to; one must also assign a value for hours. Consistent with the tendency for individuals to report usual hours ending in either a 5 or a 0, we assign the lower endpoint of the newly assigned hours bin for each individual. The results of this exercise are shown in Figure C.2b. Again we include the results from our baseline estimation shown in Figure 3 for comparison. Not surprisingly given that the earnings-hours profile remains flat, our key finding continues to hold: wages decrease significantly as hours worked increase beyond 50. In fact, the decrease is somewhat larger on account of moving individuals from the 50-54 hours bin into the 45-49 hours bin.

The above results did not stratify by age or education. Figure Figure C.3 shows that there is little variation in average weekly time use hours by age and education, and there is even less in average usual hours (not shown here).

We conclude that systematic overreporting of usual weekly hours is not the dominant explanation for the empirical pattern in Figure 3 of the main text.

Figure C.3: Average Weekly Time Use Hours



D Additional Model Results

D.1 Additional Moments for Wages and Hours

This section displays additional comparisons between outcomes in the estimated M3 model and the data. Figures D.1 and D.2 display outcomes for High School, Bachelor, and Bachelor+ men ages 50-54. See the main text for additional details on the estimation exercise.

Figure D.1: Additional Comparisons Between Model M3 and Data: High School

(a) S.D. log Wages By Hours Bin (b) Mean Hours By Wage Decile (c) S.D. Hours By Wage Decile

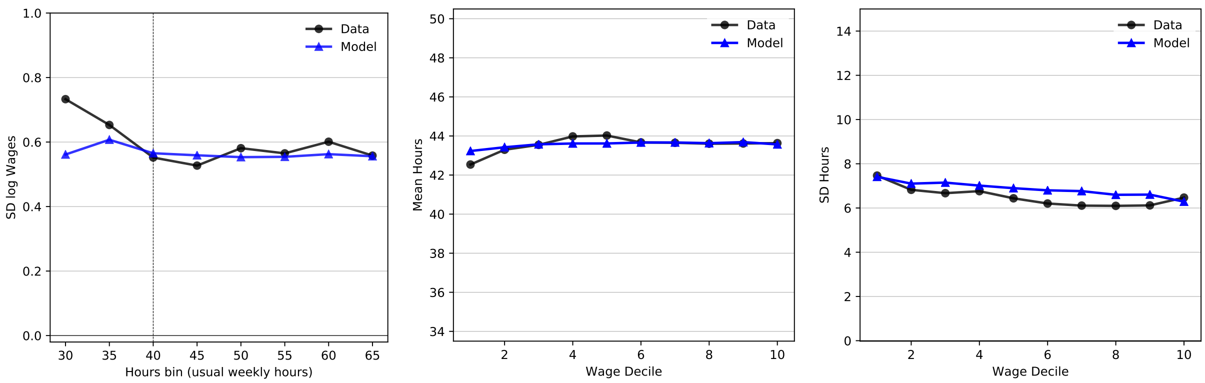
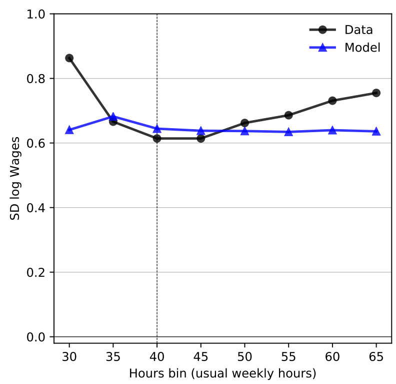


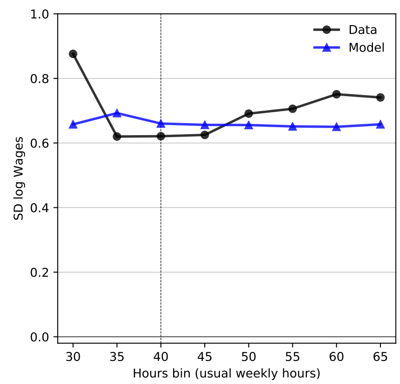
Figure D.2: Additional Comparisons Between Model M3 and Data: Bachelor and Bachelor+

Panel I: S.D. log Wages By Hours Bin

(a) Bachelor

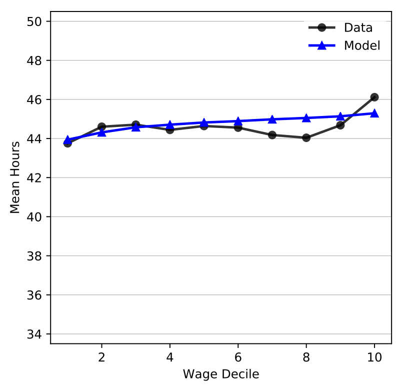


(b) Bachelor+

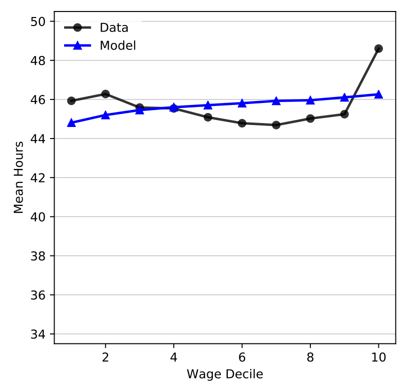


Panel II: Mean Hours By Wage Decile

(c) Bachelor

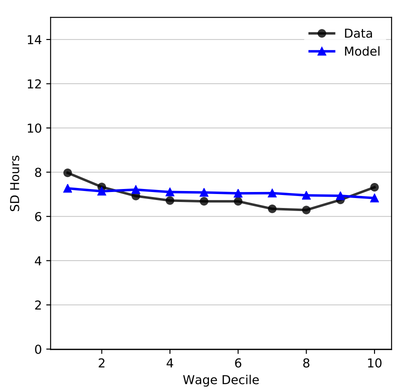


(d) Bachelor+

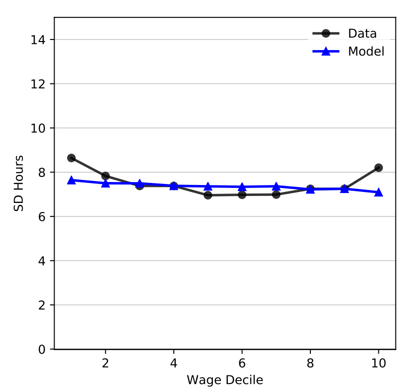


Panel III: S.D. Hours By Wage Decile

(e) Bachelor



(f) Bachelor+



D.2 Estimation of Model with Non-Labor Income

This section contains the results of our estimation exercise for High School men age 50-54 with non-labor income included. As discussed in Section 6.1, we assume that non-labor income y_i is distributed lognormally with mean μ_y and standard deviation σ_y, which we set using the SCF. We assume that y_i is uncorrelated with both α_i and z_i. We then run our standard estimation exercise.

The estimates with and without non-labor income for the 3-Region Model (M3) are displayed in Table D.1. With non-labor income, μ_α is lower to compensate for the income effect of non-labor income, and σ_α is slightly lower because there is an additional source of heterogeneity in motivations to work. With non-labor income $\rho_{\alpha,z}$ is less negative, implying a weaker positive selection of productivity across the hours distribution. This is intuitive, as non-labor income itself introduces some positive selection across the hours distribution: all else equal, a given transfer of non-labor income will reduce hours more for a low productivity worker relative to a high productivity individual.

Despite these differences, the key takeaway from Table D.1 is that the estimated earnings technology is quite similar with and without non-labor income. Specifically, $\theta_m = 0.061$ with non-labor income versus 0.058 without, and $\theta_l = 0.034$ with non-labor income versus 0.038 without. Figure D.3 shows that the overall fit of the model to the data is also barely affected by the introduction of non-labor income.

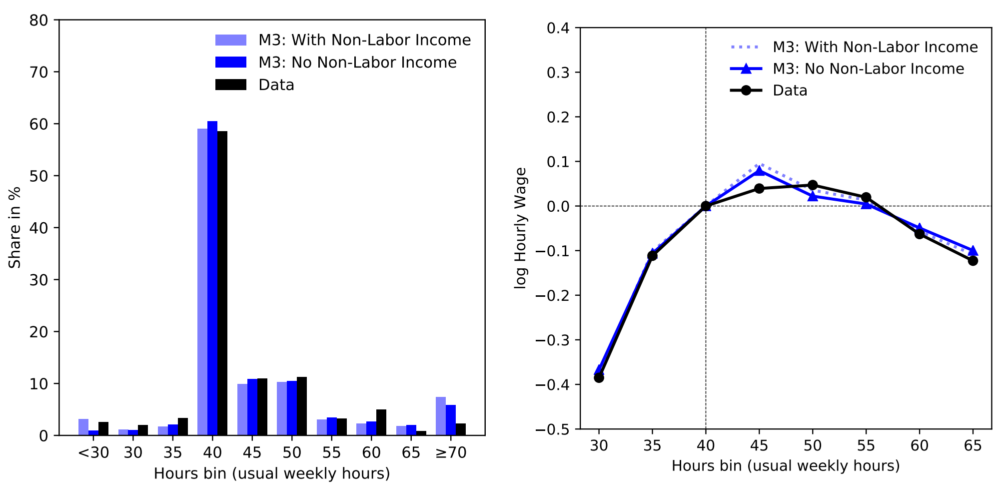
Table D.1: Estimated Parameters 3-Region Model (M3) with Non-Labor Income: High School Males 50-54

Specification	μ_α	σ_α	σ_z	$\rho_{\alpha,z}$	σ_m	θ_s	θ_m	θ_l
Benchmark	-13.64	1.61	0.589	-0.33	0.04	1.40	0.058	0.034
With Non-Labor Income	-14.22	1.50	0.583	-0.24	0.04	1.40	0.061	0.038

Figure D.3: Model Fit of 3-Region Model (M3) with Non-Labor Income: High School Males 50-54

(a) Hours Distribution

(b) Wage-Hours Profile



E Alternative Shock Process for Incomplete Markets Exercise

In Section 9.1 in the main text, we demonstrate that, within a standard incomplete markets economy, our estimated earnings technology has first order effects on the role of hours as a form of insurance. That analysis assumed that productivity followed an AR(1) process, in order to maintain comparability with benchmark results in Pijoan-Mas (2006). However, it is also standard practice to assume a shock process featuring both transitory and persistent components. In this Appendix section, we verify that such an alternative shock process does not meaningfully change our results.

We assume that assets must be non-negative, $a_t \geq 0$. We model $\log z_t$ as the sum of two orthogonal components: a persistent AR(1) process and a transitory shock:

$$\log z_t = \eta_t + \zeta_t \ ,$$
$$\log \eta_t = \rho_\eta \eta_{t-1} + \varepsilon_t \ ,$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon)$ and $\zeta_t \sim N(0, \sigma_\zeta)$.

To set parameters, we normalize w to unity and adopt the following parameterization. Following Pijoan-Mas (2006), we assume: $\sigma = 0.69$, $\gamma = 0.50$, $\beta = .94$, and $r = .05$. For the shock process we use parameter values from Heathcote, Storesletten, and Violante (2010): $\rho_\eta = 0.973$, $\sigma_\varepsilon = 0.021$, $\sigma_\zeta = 0.063$. We approximate the process for $\log z$ using the Rouwenhorst method, using 7 and 2 points for the persistent and transitory grids, respectively. We choose the value of α so as to target average hours in the ergodic distribution. As noted above, we consider three different targets for average hours: 30, 40, and 50. The rationale for these three values is that they correspond to different regions in the non-linear earnings specification: a region with convex earnings (30 hours), a region with concave earnings (50 hours) and a point at which earnings have a kink (40 hours). Note that the values of α will differ across the linear and non-linear specifications.

We then conduct the same analysis as in the main text. The resulting changes the standard deviation of hours and in welfare are displayed in Table E.1. Comparing these results to those in the corresponding table in Section 9.1, one can see that the results change little due to the different shock process.

Table E.1: Effects of Endogenizing Hours

	mean h	std h	CEV
linear earnings	40	0.22	3.8%
non-linear earnings	40	0.02	0.0%
non-linear earnings	30	0.24	6.2%
non-linear earnings	50	0.17	0.1%
non-linear earnings	60	0.16	0.3%