Hours and Wages

Alexander Bick
Arizona State University

Adam Blandin
Virginia Commonwealth University

Richard Rogerson
Princeton University

November 2018*

PRELIMINARY

Abstract

This paper examines the potential effect of hours worked on wages, both contemporaneously and dynamically. We first present an array of evidence on the relationship between wages and usual weekly hours using both cross-section and panel data. In the cross-section we document how this relationship varies with age, gender and occupation. A robust finding is that the relationship between hours and wages differs across the hours worked distribution. We next explore the extent to which a benchmark structural model of labor supply can be used to infer the effect of hours on wages. Imposing functional forms on the nature of heterogeneity we find that we can identify the profile of effects of hours on wages up to a rotation.

*Email: alexander.bick@asu.edu, ajblandin@vcu.edu, rdr@princeton.edu. Minju Jeong provided excellent research assistance.

1 Introduction

Understanding the effect of wages on desired hours of work is a classic issue in labor economics. But it has also long been understood that hours of work may affect wages.[1] Recent work on this topic has highlighted the possibility that this effect is heterogeneous across occupations and that it may play an important role in shaping some dimensions of inequality. For example, Goldin (2014) and others have argued that the heterogeneous effect of hours on wages across occupations is an important element in understanding wage inequality by gender.[2] Autor and coauthors (see, for example Acemoglu and Autor, 2011) have emphasized that occupations play a key role in shaping overall wage and earnings inequality. But if occupational wage differences partly reflect differences in hours of work then our interpretation of the occupational contribution to overall wage inequality is affected. More generally, allowing for hours of work to affect wages impacts labor supply elasticities and has implications for taxes.

The objective of this paper is to provide a more detailed examination of the effect of hours of work on wages. The starting point for our analysis is an extensive presentation of what we view as the salient facts for understanding the effect of hours on wages. Available data sources present us with a tradeoff. If one wants to consider the effects operating at the level of somewhat narrowly defined subgroups–say an age, gender, occupation cell–one needs a data set with a very large sample size. But if one wants to study dynamic effects then one needs a relatively long panel data set. An unfortunate reality is that large data sets tend to have little if any panel component, and that long panel data sets tend to have relatively few observations. In view of this situation we opt to present cross-sectional patterns using data from the 2000 Census and to present dynamic patterns using the NLSY79.

The two key objects of interest to us in the cross-section analysis are the distribution of individuals across weekly hours of work and the mean wage profile across the weekly hours distribution. For the overall population, we find a single peaked wage-hours profile, with mean wages increasing below 40 hours and decreasing above 50 hours. This basic pattern holds when we split the sample by gender, age and education. A simple message is that the cross-sectional empirical relationship between wages and hours is not well described by a single elasticity.

[1] Early examples of empirical papers that considered contemporaneous effects of hours on wages include Rosen (1976) and Moffitt (1984). Heckman (1976) is an early example of an analysis of labor supply that incorporates dynamic effects of hours on wages. See also Shaw (1989).

[2] See for example, the papers by Bertrand et al. (2010), Cortes and Pan (2016b; 2016a), and Erosa et al. (2017). See Bertrand (2017) for a survey.

When we study this relationship at the occupation level we find substantial heterogeneity across occupations, both in the shape of the wage-hours profile as well as the distribution of workers across the hours distribution. Importantly, it remains true at the occupation level that the slope of the wage-hours profile varies across the hours distribution, so that the relationship is not in general well approximated by a single elasticity as imposed by the existing literature. Moreover, the manner in which the slope changes across the hours distribution also varies substantially across occupations.

The same cross-sectional patterns for the overall population that we document in the Cenus data also hold in the NLSY79. But we can also use the NLSY79 to document the relationship between past hours of work and current wages. Importantly, we find that the correlation between past hours and current wages is significantly affected by whether we control for current hours. When we do not control for current hours there is a negative relationship between past hours and current wages. But when we control for current hours, we find that working long hours in the past is positively related to current wages. Consistent with the purely cross-sectional analysis, we find that long hours continue to be negatively related to current hours when we control for past hours. This suggests that any attempt to infer the dynamic effect of hours on wages must confront the cross-sectional relationship between hours and wages.

Having documented some salient patterns in the data, the second part of our paper explores the extent to which a simple structural model of labor supply by heterogeneous individuals can be combined with the patterns we document in cross-sectional data to make inferences about the effect of hours worked on wages, and in particular how it varies across the hours distribution. While this analysis is purely static, we see it as a key building block for future dynamic analysis.

Even subject to imposing a common structure on preferences and restricting the sources of heterogeneity for productivity and for leisure preferences to be jointly log normally distributed we find that cross-sectional data alone cannot identify the effect of hours on wages. However, we find that we can identify the shape of the effect up to a rotation that pivots around the 40-44 hours worked bin. It follows that if we had only one data point about the effect of hours on wages relative to the 40-44 hours bin we would be able to identify the entire profile of effects. We argue that quasi-experimental evidence of the sort provided by Aaronson and French (2004) could in principle be used in conjunction with our method to infer the entire slope. Extending this argument to the occupational level would require this sort of quasi-experimental evidence at the occupational level.

Relative to the existing literature, our analysis offers three key contributions. First, we allow for the effect of hours on wages to vary with the level of hours worked. That is, we do not assume that the effect of moving from part-time to full time, say from 30 to 40 hours per week, is the same as moving from 40 hours per week to 50 hours per week. While a large literature has studied the extent to which there is a wage penalty for part-time work (see, for example Aaronson and French, 2004), there is relatively little work examining the effect of moving from 40 or 50 hours per week to higher levels.

Second, we consider both static and dynamic effects. Static effects capture the extent to which working different numbers of hours today impacts the wage today. An example would be a wage penalty for part-time work. Dynamic effects capture the extent to which the choice of hours today affects future wages. Learning-by-doing would generate these types of effects. Distinguishing between the two has important implications for how differences in hours of work contribute to the overall level of inequality and how it changes over the life cycle. Imai and Keane (2004) is an important contribution to estimating the dynamic returns to current hours of work. But as we discuss in more detail below, they assume that there is no contemporaneous effect of hours of work on wages. We argue that this is likely to bias their estimates.

Third, we explicitly recognize that the cross-sectional relationship between hours of work and wages reflects not only the underlying structural effect of hours on wages but also the labor supply effects of wages on hours of work, unobserved heterogeneity and its correlation with observables, and measurement error. The recent literature has effectively run regressions of earnings on weekly hours across occupations and interpreted the coefficient as evidence of the effect of hours on wages. But absent a theory that explicitly incorporates all of the relevant elements this interpretation is unwarranted. We illustrate how inference from simple regressions can lead to mistaken conclusions and set up a theory that allows us to understand the various factors that shape the cross-sectional relationship between hours and wages. We explore the extent to which one can isolate the contemporaneous effect of hours on wages using data on wages and hours.

An outline of the paper is as follows. The next section reports some calculations in the context of a static labor supply model featuring heterogeneous households in order to illustrate some of the effects generated by models in which hours affect wages, how these effects depend on the profile of these effects across the hours distribution, and to suggest the challenges in uncovering them. Section 3 describes the data that we will use for our empirical work. Section 4 documents a set of key facts based on cross-section data, and Section 5 documents a key set of facts based on panel data. Section 6 describes a method for structurally isolating the contemporaneous effects of hours on wages using cross-section data. Section 7 concludes.

2 Motivating Calculations

Before proceeding with our analysis of patterns in the data we think it is helpful to highlight a few basic points in the context of a simple static benchmark model of labor supply among heterogeneous agents. These points serve to both motivate some aspects of the analysis and illustrate some challenges for empirical work. The first point that we illustrate is that the effect of hours heterogeneity on wage and earnings inequality depends very much on the extent to which the return to working additional hours is asymmetric across the hours distribution, as well as on the correlation between observables and unobservables.[3] Second, we emphasize that a simple cross-section regression of earnings or hours on wages is not able to properly identify the causal effect of hours on wages. Although the simple model presented here does not distinguish between contemporaneous and dynamic effects of hours on wages we nonetheless, we feel that it serves to highlight these two key messages in a very transparent way.

2.1 Model

There is a continuum of individuals of unit mass, indexed by i. Individual i has preferences over consumption (c) and hours of work (h) given by:

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

We restrict σ and γ to be the same for all households, but importantly allow for heterogeneity in preferences for consumption versus leisure as captured by the heterogeneity in α. Each individual is characterized by a productivity level, denoted by z_i. We assume that α_i and z_i are jointly log normally distributed. The standard deviation of $\log \alpha$ is σ_α, the standard deviation of $\log z$ is σ_z and the correlation between the two is denoted by $\rho_{\alpha z}$. The relative size of income and substitution effects is not restricted by this class of utility functions as this will depend on the value of σ. If $\sigma > 1$ then the substitution effect dominates, whereas if $\sigma < 1$ then the income effect dominates. In the limiting case as σ tends to one the two effects are perfectly offsetting.

Consistent with our focus on the possibility of a wage-hours schedule, we allow for the mapping between hours of work and efficiency units of labor to be non-linear, with the extent of the non-linearity varying with the level of hours worked. This mapping can be quite general, but for present purposes we assume that it

[3] In this section we abstract from differences across occupations. If the effect of hours on wages is uniform across the hours distribution for each occupation but different across occupations then this would also impact inequality.

varies across three regions of the hours distribution. Region 1 is defined by the interval $[0,\bar{h}_1]$, Region 2 is defined by the interval $[\bar{h}_1,\bar{h}_2]$, and Region 3 is the set of values for h greater than $\bar{h}_2$. In what follows we will think of Region 1 representing "part-time" work, Region 2 representing "normal" hours and Region 3 representing "long" hours.[4]

Within region j, the mapping from hours of work into efficiency units for an individual with productivity z is given by:

$$e = zA_j h^{\theta_j}$$

We normalize A_1 to equal unity and choose A_2 and A_3 so that this function is continuous across the boundaries of the three regions. Letting w be the wage rate per efficiency unit of labor, the observed wage for an individual with individual productivity z that supplies hours h that lie within Region j will be:

$$w = zA_j h^{\theta_j - 1}$$

A value of θ_j greater than unity indicates that within region j there is a positive effect of hours on wages, while a value of θ_j less than unity indicates that within region j there is a negative effect of hours on wages.

We normalize the price of an efficiency unit of labor to equal unity and also assume that consumption is measured in units such that the price of consumption is equal to unity. It follows that the budget constraint of the individual will reduce to $c = e$.

2.2 Effect of the θ_j on Inequality

In this section we use the above model to quantitatively illustrate some key points about hours of work heterogeneity and inequality. We focus on the specification in which income and substitution effects are offsetting, i.e., the limiting case in which σ tends to one and we set $\gamma = 0.25$. For the benchmark we assume that there is no causal effect of hours on wages, i.e., $\theta_j = 1$ for all θ, and set $\rho_{\alpha z} = 0$. We set the mean of $\log z$ to be zero as a normalization and set the mean level of log α such that an individual with mean values for both of z and α chooses hours equal to 40. The two remaining parameters are the standard deviations of the two idiosyncratic shocks, σ_z and σ_α. We set these equal to 0.70 and 1.25 so as to match empirical

[4]Having θ change discontinuously is potentially problematic in that it creates bunching at the points of discontinuity. To the extent that we see bunching around 40 hours this may be an empirically relevant feature. But we do not see the same degree of bunching at any higher level of hours, suggesting that a more appropriate specification might be one in which θ changes continuously at hours of work above 40. Here we report results based on the simpler three region specification to facilitate exposition, but later on in the paper we will allow theta to change continuously.

Table 1: Effect of θ_j's on Inequality $\rho_{\alpha z} = 0.00$

θ_1	θ_2	θ_3	sd(log w)	sd(log e)	sd(log h)
1.00	1.00	1.00	0.70	0.74	0.25
1.40	1.40	1.40	0.71	0.78	0.25
1.00	1.00	1.40	0.77	0.93	0.30
1.40	1.00	1.00	0.70	0.75	0.23

counterparts for the cross-sectional standard deviation of log wages and log usual weekly hours in the data.

In what follows we report results from simulating a sample with $100,000$ individuals. In this benchmark setting, the standard deviation of observed log wages, i.e., the log of earnings divided by hours, is identical to the standard deviation of $\log z$ and so is also equal to 0.70. The standard deviation of log hours is 0.25 and the standard deviation of log earnings is 0.74. These results are shown in the first row of Table 1.

We now consider the effect of allowing for a causal effect of hours on wages. In these exercises we keep all of the other parameters constant as we change the values of the θ_j except for the mean of log α which is always adjusted so that the individual with mean values for both $\log z$ and $\log \alpha$ chooses 40 hours. Our main goal is to illustrate the extent to which changes in the θ_j can impact the level of observed wage inequality relative to the underlying dispersion in productivity.

For our first exercise we assume that the θ_j are all equal but greater than unity. For purposes of illustration, we adopt the estimate of θ from Aaronson and French (2004) and set $\theta_j = 1.4$ for all j. For this specification we find that the standard deviations of log wages, log earnings and log hours are now equal to 0.71, 0.78, and 0.25 respectively, implying that there is only a very modest increase in the level of wage inequality. Recall that we are keeping the standard deviation of $\log z$ fixed at 0.70.

For our next exercise we assume that the positive effect of hours on wages is present only for individuals working long hours, i.e., we assume that $\theta_1 = \theta_2 = 1$ and $\theta_3 = 1.4$. For this specification we find that the standard deviations of log wages, log earnings and log hours are now equal to 0.77, 0.93, and 0.30 respectively. Importantly, this asymmetry in the return to working additional hours has much larger effects on wage inequality and earnings inequality. It also induces significant effects on hours inequality with more people shifting to the upper end of the hours distribution. When θ was above one and constant, it created an incentive for all individuals to increase their hours of work. But when θ is only above one for individuals working long hours there will be no effect on hours for people who are working below some threshold, and it is only those working long hours that are incentivized to work additional hours. This explains why the dispersion in hours increases.

Next consider the case in which we assume $\theta_2 = \theta_3 = 1$ and $\theta_1 = 1.4$. For this specification we find that the standard deviations of log wages, log earnings and log hours are now equal to 0.70, 0.75, and 0.23 respectively. Notably, in this case there is essentially no increase in wage and earnings inequality, and intuitively, a modest reduction in the dispersion of hours. Although one might have expected the nonlinearity to induce added dispersion in wages, the reason this does not show up in the statistics is that the wage penalty for working part-time induces people to avoid part-time work. But importantly, even those who continue to work part-time now work more hours than they would have relative to the benchmark case in which $\theta_j = 1$ for all j, thereby dampening the overall effect.[5] Contrasting these last two cases, a key difference is that that when $\theta_1 = \theta_2 = 1$ and $\theta_3 = 1.4$, individuals gravitate toward the region in which the causal effect is positive, whereas when $\theta_2 = \theta_3 = 1$ and $\theta_1 = 1.4$, individuals gravitate away from the region in which the causal effect is positive. Related to this, we note that the case of $\theta_2 = \theta_3 = 1$ and $\theta_1 = 1.4$ produces a greater concentration of individuals around 40 hours, a feature that is prevalent in the data.

There are two key messages from these examples that we want to emphasize. First, the effect of hours heterogeneity on wage inequality depends very much on the profile of θ's across the hours distribution. In the examples above, the greatest effect occurs when the causal effect is positive only for those working long hours, and is of second order importance in all of the other cases. Second, the effect very much depends on changes in the distribution of hours worked; the large effect for the case in which $\theta_1 = \theta_2 = 1$ and $\theta_3 = 1.4$ is due at least in part to the fact that this induces a large fraction of individuals to work long hours. Looking ahead to the exercises that we carry out later in the paper, this suggests that it is important to match the hours distribution in order to infer the impact of hours on wages.

The previous benchmark specification assumed that productivity and preferences for work were uncorrelated. Given that preferences had offsetting income and substitution effects it followed that in the benchmark specification there was zero correlation between wages and hours. The next set of results considers specifications in which there are positive or negative correlations between wages and hours in the benchmark. We find that even modest correlations can substantially impact how the θ_j affect inequality. We first set $\rho_{\alpha z} = -0.25$ and repeat the same exercises as before, which implies that high productivity individuals have on average low preferences for leisure. Results are displayed in Table 2.

[5] Although not reported in the table, we can also consider the case in which $\theta_1 = \theta_3 = 1$ and $\theta_2 = 1.4$. Perhaps not surprisingly, this creates some compression in hours, and has a very modest effect on inequality.

Table 2: Effect of θ_j's on Inequality $\rho_{\alpha z} = -0.25$

θ_1	θ_2	θ_3	sd(log) w	sd(log e)	sd(log h)
1.00	1.00	1.00	0.70	0.74	0.23
1.40	1.40	1.40	0.74	0.86	0.25
1.00	1.00	1.40	0.83	1.03	0.30
1.40	1.00	1.00	0.71	0.81	0.23

It remains the case that when all of the θ_j are equal to one that the standard deviation of log wages is identical to the standard deviation of $\log z$. If we now set $\theta_j = 1.4$ for each j we find that the standard deviations of log wages, log earnings and log hours are now equal to 0.74, 0.86, and 0.25 respectively, implying that there is a modest increase in the level of wage inequality relative to the underlying dispersion in productivity. If we set $\theta_1 = \theta_2 = 1$ and $\theta_3 = 1.4$ we find that the standard deviations of log wages, log earnings and log hours are now equal to 0.83, 1.03, and 0.30 respectively. And for the case of $\theta_2 = \theta_3 = 1$ and $\theta_1 = 1.4$ we find that the standard deviations of log wages, log earnings and log hours are now equal to 0.71, 0.81, and 0.23 respectively.

Importantly, if we compare the changes as we go from the first row to the third row in Tables 1 and 2 we see that the increase in wage inequality is almost twice as large when $\rho_{\alpha z} = -0.25$. This is intuitive–when $\rho_{\alpha z}$ is negative it implies that individuals who work long hours tend to be the people who are most productive, so that increasing the wages of these individuals tends to increase dispersion in wages by increasing wages at the top of the distribution.

We have also repeated the same exercise for the case of $\rho_{\alpha z} = 0.25$, implying a modest negative cross-sectional correlation between hours and wages in the benchmark cross-section. We now find that non-linearities yield very modest changes in inequality in all cases, with a modest decrease in the case of $\theta_1 = \theta_2 = 1$ and $\theta_3 = 1.4$. This is intuitive: with $\rho_{\alpha z} = 0.25$ it is the low productivity individuals who tend to work long hours, so when $\theta_3 = 1.4$ it is these individuals who are motivated to work longer hours and hence increase their wages. The upshot is that the effect of the θ_j on inequality depends crucially on the underlying correlation between productivity and hours.

2.3 Uncovering the θ_j

Having shown the potential for the θ_j to have significant impacts on inequality, we now want to note one reason why uncovering the θ_j profile from cross-section data is difficult. The challenge is simple: what one wants to know is how the wage for a particular individual would change if that individual were to change his or her hours of work holding all else constant. But observing the wages of other observationally similar workers who work different hours need not be informative about this if we cannot control for unobservables that might be correlated with wages. Put somewhat differently, if we could observe the change in earnings as we vary hours for a given individual, and we were confident that the individual's unobserved productivity was remaining constant, then a regression of log earnings on log hours would deliver an estimate of θ. But this of course need not be the case if we run a regression of wages on hours using cross-sectional data.[6]

As a simple illustration of the quantitative magnitude of the problem, we regress log earnings on log hours using cross-section data generated by the above model. This is an exercise that several papers have followed as a way to learn about possible causal effects of hours on wages (see, for example, Goldin, 2014). For simplicity we consider the case in which all of the θ_j are equal to 1.4, i.e., we assume a constant elasitcity along the hours distribution. We use the parameters as in the calculations above, except that we consider three different values of $\rho_{\alpha z}$: $-0.25, 0$, and 0.25. The three estimated coefficients on log hours are 2.00, 1.39 and 0.67 respectively for the three values of $\rho_{\alpha z}$. It follows that using this regression coefficient to make inference about the value of θ may lead to very biased conclusions about the returns to working additional hours. In this section we have focused on the case in which σ tends to unity. If we were to consider departures from this value we would obtain similar conclusions about the connection between the regression coefficient on log hours and the true value of θ.

The key message is that one needs a more sophisticated strategy to uncover the value of θ. In particular, although the cross-sectional relationship between hours and earnings contains information about θ, it is critical to properly control for other influences in order to extract the information about θ. This is the challenge that we take up later in this paper.

A second point that we want to make is that even when there is no issue of correlation between observables and unobservables, a simple regression of wages on hours assuming a constant elasticity across the hours distribution is problematic because the elasticity may vary with hours. As indicated in our previous examples, this regression coefficient will not be a sufficient statistic for thinking about the effects of hours on wages and earnings. To illustrate this assume $\rho_{\alpha z} = 0$ and consider the three different settings for the θ_j profile studies in Table 1. As noted above, if we estimate a constant elasticity relationship for the case in

[6]Note that panel data does not immediately solve this issue if we are concerned that changes in an individual's hours are correlated with changes in that individual's unobserved productivity.

which each of the $\theta_j = 1.4$ we obtain an estimated of 0.39 when regressing log wages on log hours. If only θ_1 is equal to 1.4 then the estimate is 0.16 and if only θ_3 is 1.4 then the estimate is 0.98. The message is a simple one: if the profile of the θ_j is not constant then it is not sufficient to try to summarize the relationship between hours and wages with a single elasticity.

3 Data Sets

An important reality for work in this area is the different data requirements for distinguishing between static and dynamic effects. Analyzing dynamic effects requires panel data. But a key limitation of the available panel data sets with data on hours of work, earnings and occupation is that they have relatively small sample sizes. This limits the extent to which one can study relations at a more disaggregated level. On the other hand, relying only on cross-sectional data allows one to access either the CPS, with roughly an order of magnitude more observations than typical panel data sets, or the Census, which is even an additional order of magnitude larger, thereby permitting one to consider occupational differences at a very fine level. The downside to relying on cross-section data is that one cannot distinguish between static and dynamic effects.

Given this reality, our empirical analysis proceeds in two phases. The first phase documents basic cross-sectional patterns in earnings and weekly hours, both in the aggregate as well as disaggregated by gender, age, education and occupation. Our main data source is the 2000 Census, though we have also confirmed where feasible that the same patterns emerge from the CPS. Beyond standard controls (education, race, marital status, etc.), the variables of interest are: annual earnings last year, usual hours worked per week last year, weeks worked last year, and occupation last year.[7]

Our benchmark cross-sectional results are based on the 2000 Census and focus on individuals aged 25 to 64 with a positive sampling weight. We follow fairly standard definitions for working: having worked at least one week in the previous year and usual hours worked per week being larger or equal to 10. We exclude self-employed individuals and those who are enrolled in school. That leaves us with 2,504,866 men and 2,298,047 women.

The next phase extends the analysis to panel data from the NLSY79. The NLSY79 has several unique features, which we exploit. First, it allows us to observe workers over nearly their entire career: individuals were first interviewed between ages 14-22, and were 49-56 years old in the most recently available interview

[7]Note that for the Census we only know the current occupation; however, using the CPS, for which we have last year's and current occupation, we show that this does not affect our results.

wave. Note that the PSID, an alternatively publicly available longitudinal data set, has a much higher attrition rate and thus much fewer individuals observed for such a long time span. At age 25, the youngest age of our analysis we have 4178 men and 3897 women who satifisy our work criteria (at least 10 hours worked per week, at least worked one week), at age 49 (i.e., the age of the youngest NLSY79 cohort in the most recent wave) we still have 2719 men and 2886 women in our sample. At age 54, the oldest age we include in our analysis, we are left with 776 and 912 men. On average we observe individuals for around 22 years between ages 25 and 54. Second, the NLSY79 work history is more detailed: it allows us to track hours even during years when individuals were not interviewed.

4 Facts from Cross-Sectional Analysis

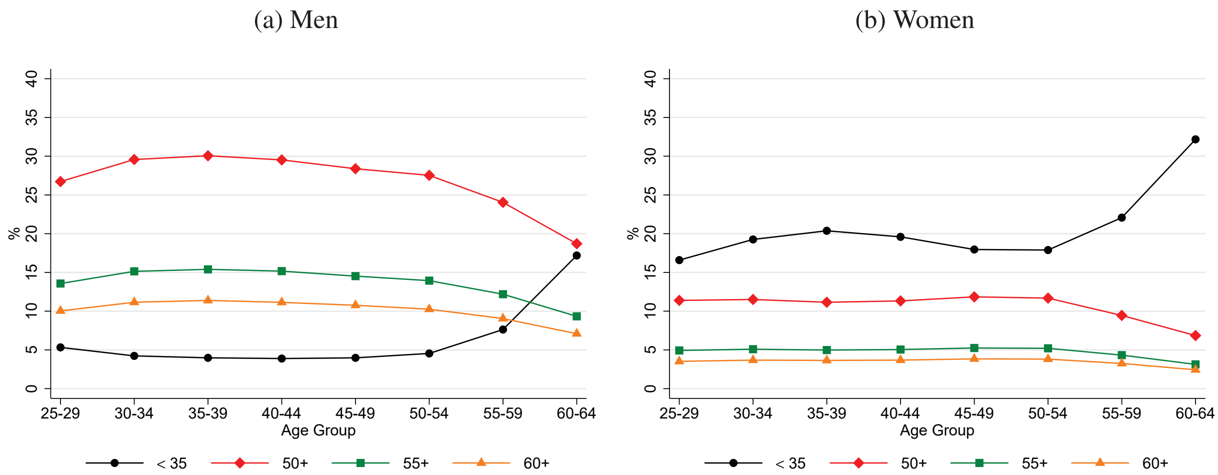
In this section we summarize a set of novel facts about the cross-sectional relationship between hours and wages that we find striking and that will serve as the focus of our subsequent structural analysis. We also examine how this relationship relates to the distribution of individuals across the hours worked distribution. We are particularly interested in how the relationship between hours and wages varies across the hours distribution, and in particular within the "long hours" region, which we will define as those working more than 50 hours per week. As noted earlier, this focus is motivated both by the work of Goldin (2014), who emphasized the potential for higher remuneration of long hours worked for the gender pay gap, as well as the simulations in the earlier subsection.

4.1 Incidence of Long and Short Hours

We start by examining the distribution of individuals across the weekly hours distribution by age, gender and education. In what follows we will label individuals who work at least 50 hours per week as working long hours, those working less than 35 as working short hours, and those in the interval 35-49 as working normal hours. In our regression analysis later on there will be no significance to these thresholds. Figure 1a shows how males in our sample are distributed across long and short hours by age.

Five features are worth noting. First, long hours are relatively common among males, with more than 25% of all males working more than 50 hours per week as their usual hours. Second, short hours are relatively uncommon, accounting for only about 5% of observations. It follows that roughly 70% of all males have usual hours in the normal range. Third, the incidence of long hours is very stable over the age

Figure 1: Fraction Working Short, Long Hours and Very Long Hours



range of 30-50. It is slightly lower in the 25-29 and the 50-59 age ranges, and much lower in the 60-64 age range. Fourth, while a large share of those working long hours are working 55 or fewer hours, roughly 15% of prime aged males work more than 55 hours and roughly 10% work more than 60 hours per week. Fifth, short hours are relatively stable across the life cycle except for a significant increase for those above 60.

Figure 1b shows the comparable figure for women. Relative to the patterns for males, two points stand out. First, long hours are much less prevalent, averaging around 10% for females versus 25% for males. Second, short hours are much more prevalent, averaging around 20% for females versus 5% for males. It follows that for both groups roughly 70% of workers have hours in the normal range. Beyond these level differences, the other patterns are basically the same: stable incidence of long and short hours for prime aged women and a sharp increase in short hours for those aged 60-64.

Figure 2 shows the incidence of long hours by education, gender and age. There is a strong correlation between the incidence of long hours and educational attainment: for prime aged males with more than a college degree we see that more than 40% or them work long hours. However, it is also noteworthy that even among males with less than high school, there are still roughly 20% of them who work long hours. A similar pattern holds for women, though the levels are all lower. Interestingly, the incidence of short hours varies much less by educational attainment. This is shown in Figure 3.

It is also of interest to examine the variation of the incidence of long and short hours by occupation. For this we consider occupations at the 3 digit level, and for each occupation we calculate the incidence of long and short hours, separately by gender. Figure 4 shows the scatter plots for males and females, with each

Figure 2: Fraction Working Long Hours by Education

(a) Men (b) Women

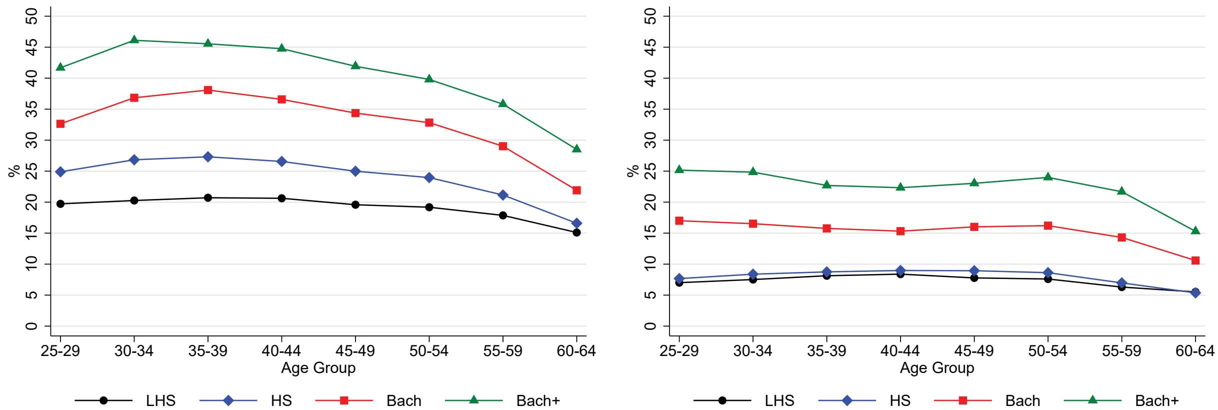
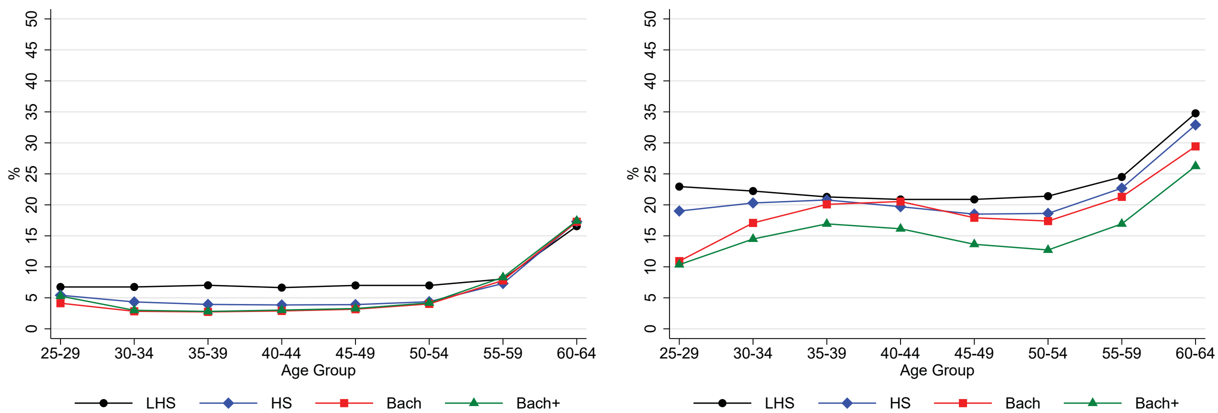


Figure 3: Fraction Working Short Hours by Education

(a) Men (b) Women

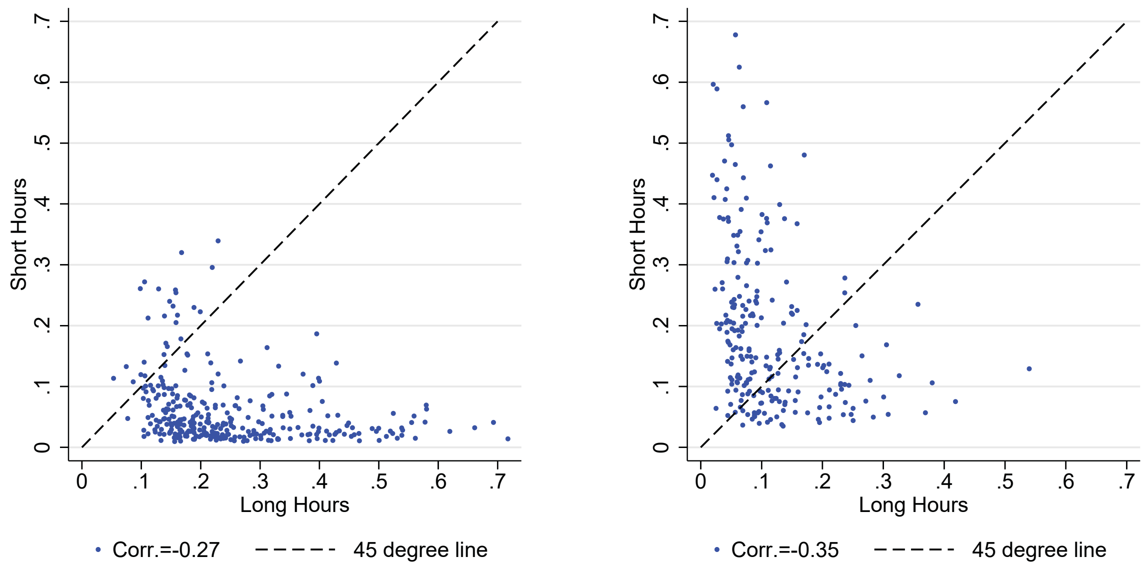


point representing the combination of long and short hours incidence for a given occupation.

A few patterns are worth noting. First, note that there is significant variation in the incidence of long and short hours across occupations. For males the range of long hours incidence runs from 5% to over 70%, with the prevalence of short hours ranging from virtually 0% in many occupations up to 35%. For women there are multiple occupations with shares of short- and long-hour workers below 10%. The share of female short-hour workers goes up to 70% and of female long-hour workers up to 55%. Second, the correlation between long and short hours across occupations is negative for both sexes, -0.27 for males and -0.35 for females. Hence, occupations with a higher share of short hours workers are generally occupations with a

Figure 4: Prevalence of Long vs. Short Hours by Occupation

(a) Men (b) Women



Unit of observation is a 3-digit occupation in the 2000 Census. Sample restricted to occupations with at least 1,000 male observations and 1,000 female observations. This reduces the sample size from 451 occupations to 197 occupations. Although we lose a large number of occupations, we note that most of these occupations are relatively unimportant in terms of employment and so we only lose about 5

lower share of long hours workers.[8] We can also compute the extent to which the relative incidence of long and short hours across occupations is gender specific. In particular, we find that the correlation of male and female incidence of long hours across occupations is 0.93, with the corresponding figure for short hours being 0.90.

4.2 Wages and Hours

Motivated by our earlier simulations, we want to characterize the relation between usual weekly hours and wages that allows for the nature of this relationship to vary across the hours distribution. This leads us to generalize the analysis in Goldin (2014) and run the following regression on cross-section data:

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

[8] The fact that the shares of workers must sum to one would mechanically generate a large negative correlation if all of the adjustment involved short and long hours. The reported correlations suggest that a lot of the adjustment involves movement between normal hours and the other two categories.

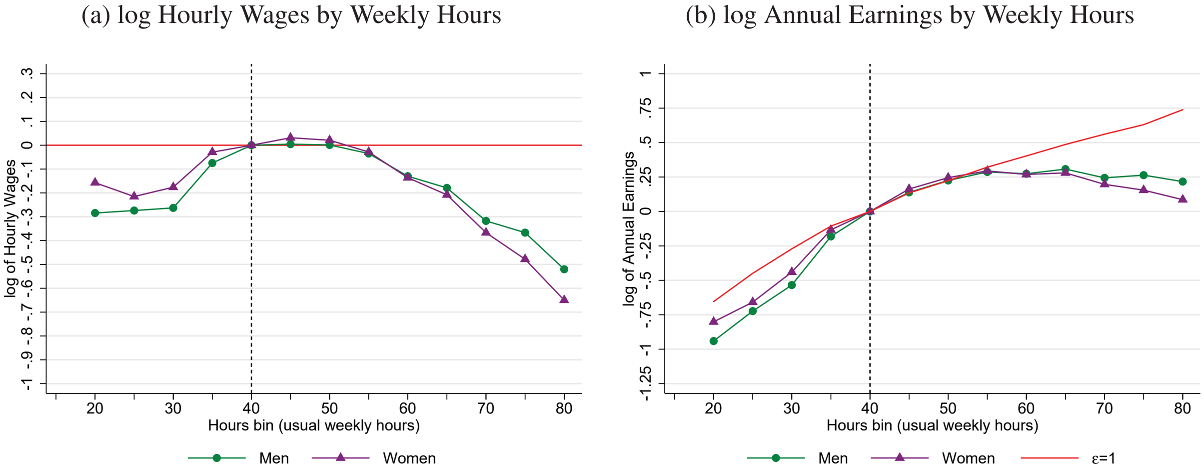
where i denotes an individual in the 2000 Census, w denotes log wages, i.e. the log difference between annual earnings and annual hours worked in 1999, wks denotes log weeks worked in 1999, and X is a vector of controls including a quadratic in age, three educational dummies (less than high school, bachelor's degree, and graduate degree), a dummy for whether the individual is married, and dummies for if the individual is black or Hispanic. Our hours measures refers to usual weekly hours worked in 1999. Importantly, we do not attach any causal significance to this estimated relationship–in the next section we outline a strategy to isolate the underlying causal relationship by using these estimates as moments to be matched in the context of a structural model of labor supply.

The key difference relative to Goldin (2014) is that instead of simply including log weekly hours in the previous year as a right hand side regressor, which imposes a constant elasticity relationship between wages and hours, we allow this elasticity to vary with hours non-parametrically.Our approach is to partition the range of weekly hours (from a minimum of 10 to a maximum of 99) into a set of 5-hour bins: 10-14, 15-19, ..., 75-79, 80-99 (the final bin runs from 80-99 hours because there are few observations above 80 hours). We denote the set of hours bins by $H = \{10, 15, ..., 80\}$, where $h \in H$ denotes the minimum threshold of a particular hour bin. The regression includes a set of dummies $\mathbb{1}_{ih}$ which equal one whenever i's weekly hours lies in bin h. The coefficients of interest are the corresponding β_h. Note that if the elasticity of wages with respect to weekly hours would be indeed independent of the hours level, this would be picked up accordingly by the coefficients for the hours dummies. This regression can be run on subgroups defined by age, education, gender and occupation. For illustrative purposes, we start by showing the results by gender, grouping all ages, education groups and occupations together. Figure 5a plots the estimates for β_h relative to the estimate for the 40 hours bin for our sample of men and women separately. We view the 40 hour bin as a natural reference point because roughly half our sample falls in this bin.

As hours increase from the part-time to the normal region, hourly wages grow. For example, at the 30 hours bin and below hourly wages are 15-25 log points lower than in the 40 hours bin. Next, hourly wages are roughly constant throughout the normal hours bins. As hours increase in the long hours region, however, hourly wages begin to fall. In fact, hourly wages in the 70 hours bin are lower than hourly wages in the 30 hours bin.

Many papers in the literature actually regress log annual earnings or log weekly earnings on log weekly hours (plus controls) in order to estimate the elasticity of earnings with respect to weekly hours. In this case the natural benchmark is a unitary elasticity as this would imply that earnings are proportional to weekly

Figure 5: log Hourly Wages, log Annual Earnings and Weekly Hours Worked



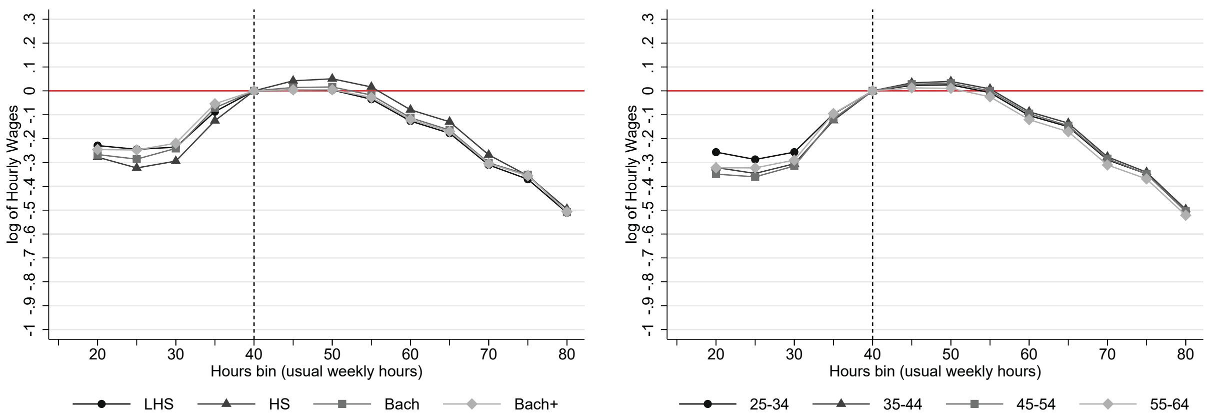
hours, or put differently the implied hourly wage would be constant. In fact, if we would use dummies for each value of weekly hours (which are only reported as integers in the data) the two regressions would yield an exact one to one mapping between the respective coefficients for hours. In Figure 5b, we run the same regression as for wages but with log annual earnings as dependent variable. This shows why the wage decreases for hours above 50: earnings are essentially flat above 50 hours.

A final takeaway from Figure 5a is that we observe the same basic patterns for both men and women. Gender differences are only present for low and very high hours, and even in those ranges the differences are fairly small. (Note that because the curve for each gender is normalized relative to the 40 hours bin, this figure does not provide information about the level of the gender gap.) The same patterns emerge if we split the sample by age group and education. Figure 6 presents the results for males. In the interests of space we only show results for men, but the message is the same for both men and women.

One potential concern is that the comparatively flat earnings coefficients above 50 hours shown in Figure 5b are driven by top-coding, but we argue that this is not of first order importance. First, even among long hours workers the fraction of top-coded men and women is low – around 4% and 2%, respectively – and this fraction varies little with hours worked (conditional on working long hours). One can assess the size of the effect of top-coding by assuming different earnings-hours elasticities across long hours bins for the top-coded. As a robustness check, we estimate such elasticities using the PSID, which features much higher top-coding thresholds. This implies only small adjustment. Second, instead of estimating OLS we estimate a quantile regression at the median. The implied estimated coefficients are much less sensitive to observations

Figure 6: log Hourly Wages and Weekly Hours Worked for Men

(a) By Education (b) By Age



at the top and the bottom than the OLS ones but are virtually identical to the OLS estimates.

Another concern is that these estimates are affected by measurement error. If people with high hours tend to be people who have overreported their hours then this will show up as a negative effect of hours on wages. For now we do not attempt to make any correction for this, but our structural analysis later on in the paper will explicitly incorporate measurement error.

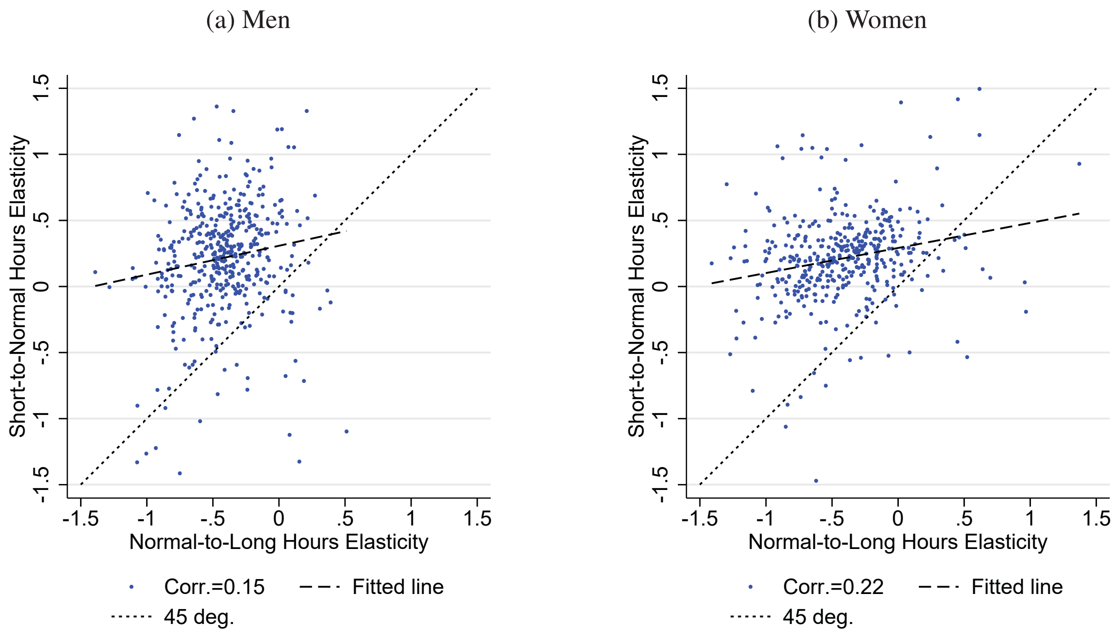
4.3 Results by Occupation

To this point we have presented results for workers from all occupations combined. One can repeat this analysis for each occupation. Below we present results for a few selected occupations which we think are of interest, but first we attempt to summarize some aspects of the distribution of effects across occupations. For each occupation we first compute mean log wages and mean log hours in three hours regions: part-time (10-34 hours), normal (35-49 hours) and long hours (50-99 hours). We then compute a "short-to-normal" and "normal-to-long" elasticity, which is the change in log mean wages divided by the change in log mean hours as one moves from the short region to the normal region and from the normal region to the long region, respectively.[9]

Figure 7 plots these two elasticities for each occupation separately by gender. Consistent with the aggregate analysis, we continue to find strong evidence of a non-linear relationship: the "short-to-normal" elasticities are usually larger than the "normal-to-long" elasticities, which in turn are for the majority of

[9]We include the same set of controls as previously when computing these elasticities.

Figure 7: “Short” and “Long” Wage-Hours Elasticities by Occupation



occupations negative. However, there is substantial heterogeneity across occupations in terms of the nature of the non-linearity and the correlation between the two elasticities is quite small. This also shows that differences in “average” elasticities across occupations can be very misleading regarding elasticities across occupations at a given point in the hours distribution. This claim is confirmed by Figure 8, which plots the two set of elasticities against the constant elasticity estimated in the literature, i.e. from regressing log wages on log hours: the correlation between this “constant” elasticity and the “short-to-normal” and “normal-to-long” elasticities is positive but far below one. For brevity, we only show here the results for men, which are however very similar to those for women.

Another striking fact is shown in Figure 9, which displays the relationship between the hours distribution and the wage-hours elasticity across occupations. Here we plot for each gender the “short-to-normal” elasticity and “normal-to-long” elasticity against the share of workers in an occupation working short and long hours, respectively.

One might expect to see more people working short hours in occupations with a low“short-to-normal”elasticity, i.e. an wage-hours elasticities between short and normal hours close to zero. Alternatively, one might expect to see more people working long hours in occupations with a higher wage-hours elasticities between normal and long hours. While that intuition is qualitatively confirmed, quantitatively the relationship it is quite

Figure 8: “Constant”, “Short”, and “Long” Earnings-Hours Elasticities by Occupation for Men

(a) Short vs. Constant

(b) Long vs. Constant

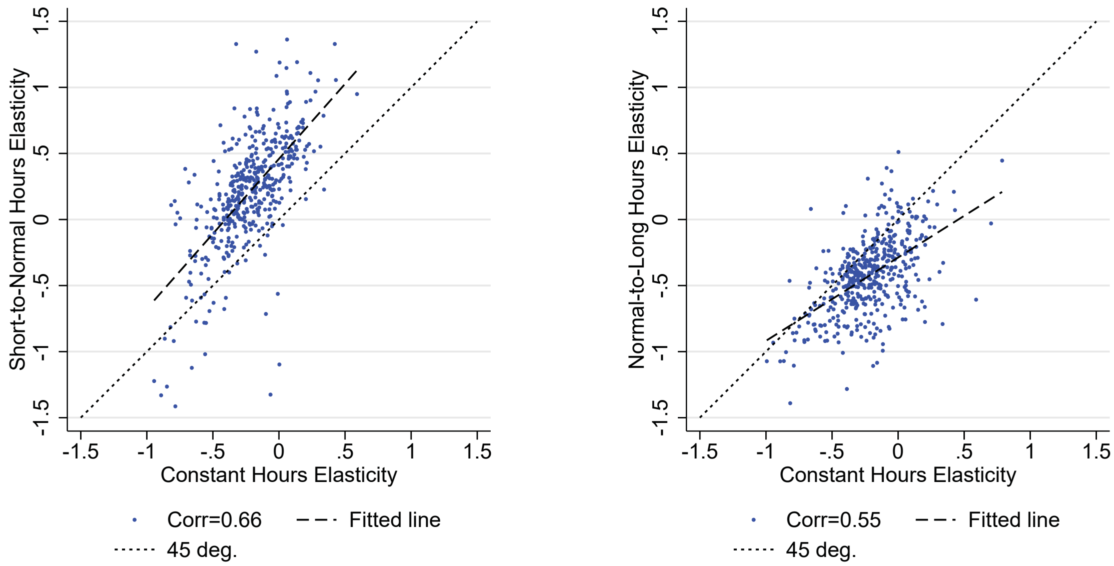


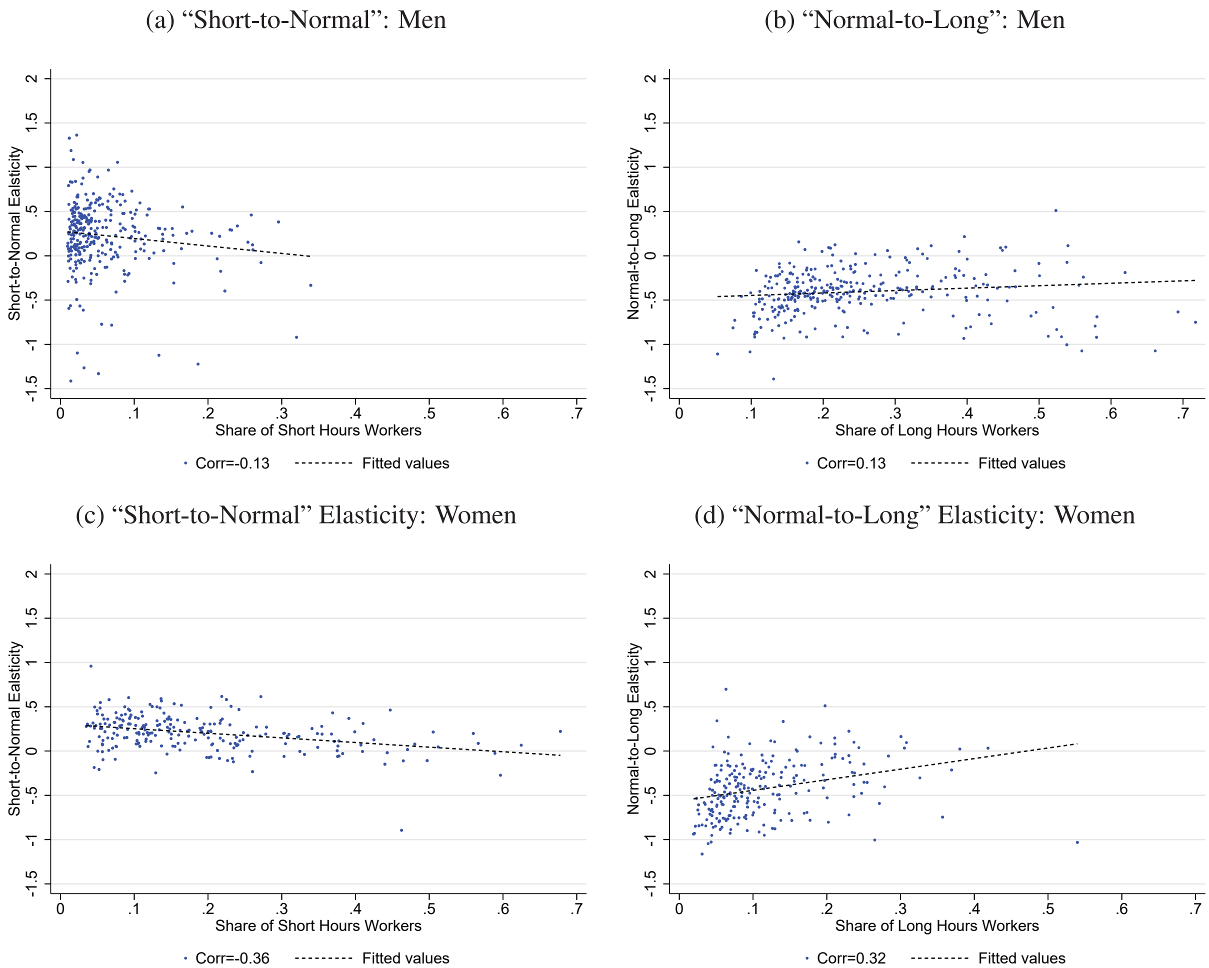
Table 3: Distribution of Workers by Hours

Fraction	Lawyers		Surgeons		Registered Nurses		Pharmacists	
	Male	Female	Male	Female	Male	Female	Male	Female
< 35	2.7	10.4	4.0	12.5	7.6	25.8	5.1	24.5
35-49	42.7	51.0	26.2	32.5	75.0	64.8	75.9	66.4
50-60	48.5	34.6	43.8	33.4	14.3	7.9	16.5	7.9
> 60	6.1	4.0	26.1	21.6	3.2	1.4	2.5	1.2

weak, especially for men.

For future reference we also present the results for four specific occupations: lawyers, pharmacists, registered nurses, and physicians and surgeons, separately for males and females. Three of these occupations were explicitly discussed in Goldin (2014): lawyers and physicians were examples of “high-powered” occupations with high mean hours, while pharmacists were a somewhat atypical example of an occupation with (i) high skill requirements and high mean hourly wages, (ii) modest mean hours, and (iii) a relatively small gender wage gap. We also include registered nurses; interestingly, we find that they have a similar hours-wage profile to physicians and surgeons, but a very different hours distribution. Table 3 presents results on the share of male and female workers working different usual hours for each of the four occupations and Figure 10 shows the hours-wage profiles for each occupation for males and females.

Figure 9: Wage-Hours Elasticities and Hours Distribution

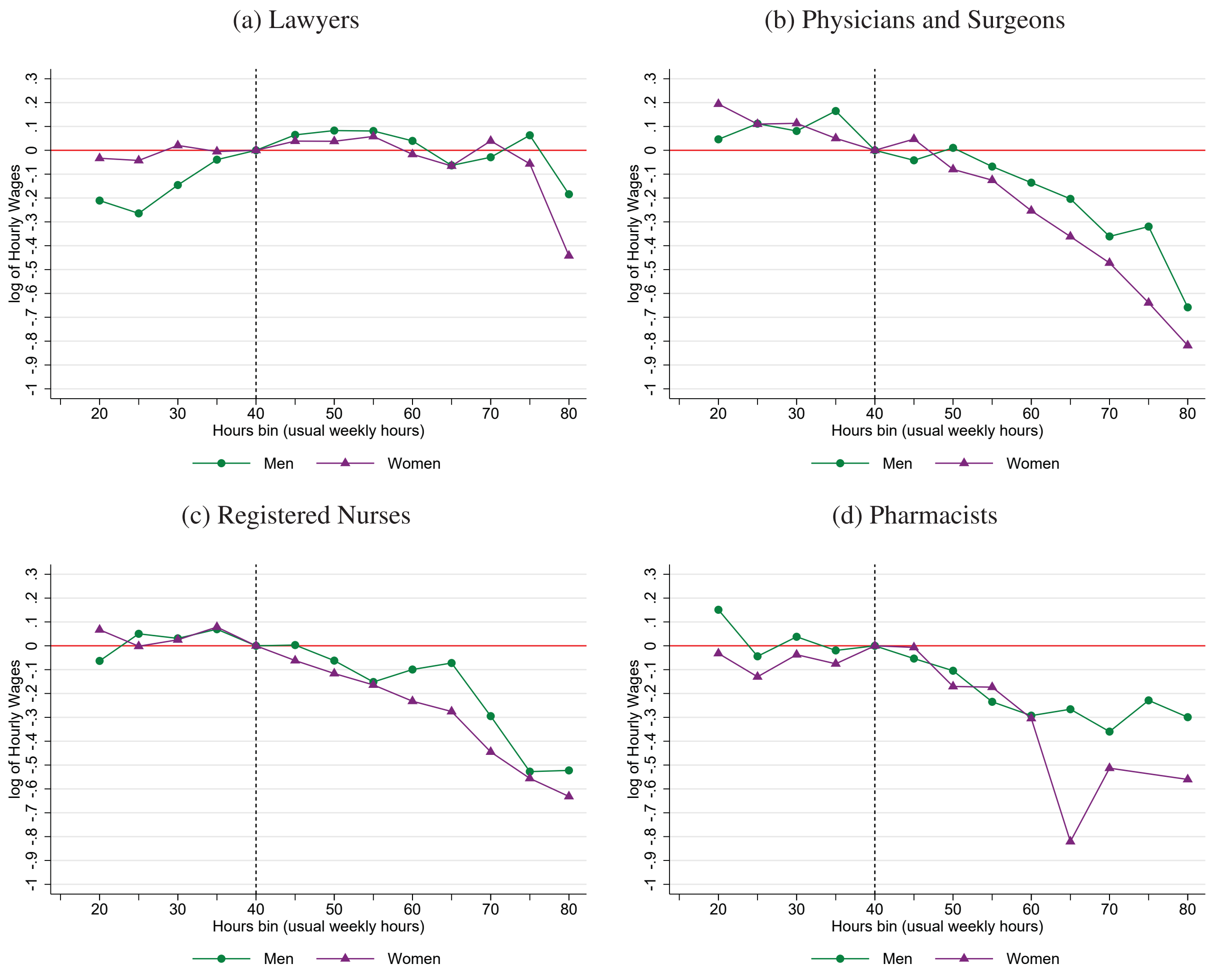


At this point we want to emphasize that even among these four occupations there is a lot of variation in terms of the distribution of hours worked and and the shape of the wage hours profiles. We think it is interesting that the wage-hours profiles for surgeons and physicians, registered nurses and pharmacists have the same qualitative shape while the hours distributions for the latter two occupations are dramatically different than for surgeons and physicians.[10] We will return to this later. Another interesting aspect is (as for the aggregate) the minimial variation between the wage-hours relationship by age, see Figure 11. Again, for brevity we only show the results for men but they are essentially identically to those for women.

Because some of these occupations are relatively high paying it is relevant to revisit the issue of top-coding and how it might affect the patterns that we find. One might conjecture that the reason that wages

[10]Although three of these occupations have hours-wage profiles that are somewhat close to linear, this is not at all the typical pattern, as our earlier figure on short and long hours elasticities indicated.

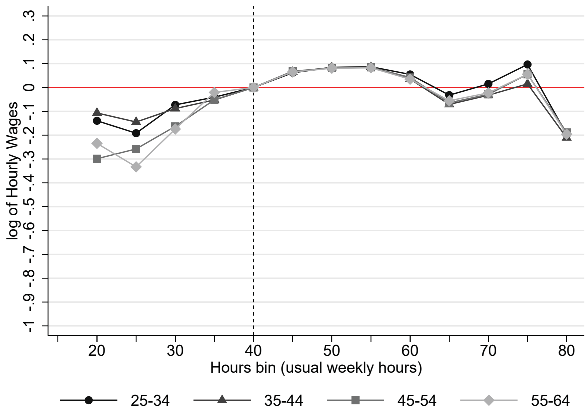
Figure 10: Wages and Hours by Occupation



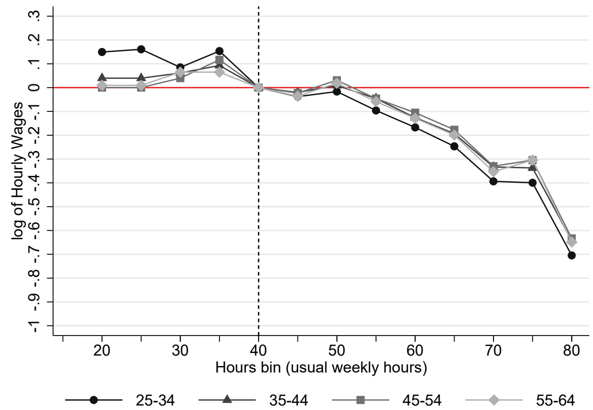
decrease with hours for surgeons, for example, is that the incidence of top-coding increases with hours, thereby artificially reducing mean wages since top-coded individuals are assigned the same level of earnings independently of their hours of work. In fact, the incidence of top-coding for this group is significant and does increase with hours worked. Figure 12 shows the fraction of top-coded individuals by hours bin for lawyers as well as physicians and surgeons (top-coding is rare for pharmacists and registered nurses even for high hours worked). Importantly, even for these occupations, the incidence of top-coding in the higher hours bins remains less than 50%, which means that if we were to instead focus on median wages by hours bin top-coding would have no impact. We mentioned earlier for the overall sample that the mean and median tracked each other quite closely. This remains true even for specific occupations. Figure 13 shows the wage-hours profile for both lawyers and surgeons based on both OLS (mean) and a quantile regression run at the median. This leads us to believe that top-coding is not the driving force for the negative slope for longer

Figure 11: Wages and Hours by Occupation and Age for Men

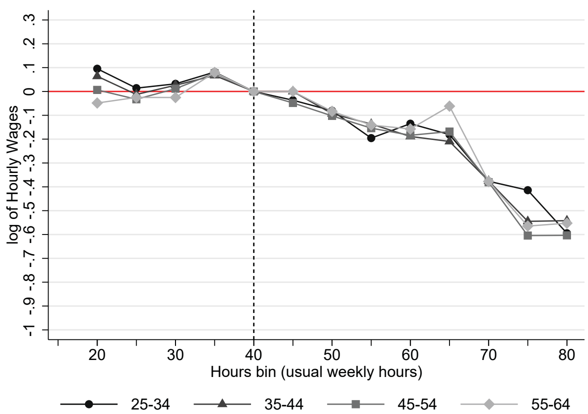
(a) Lawyers



(b) Physicians and Surgeons



(c) Registered Nurses



(d) Pharmacists

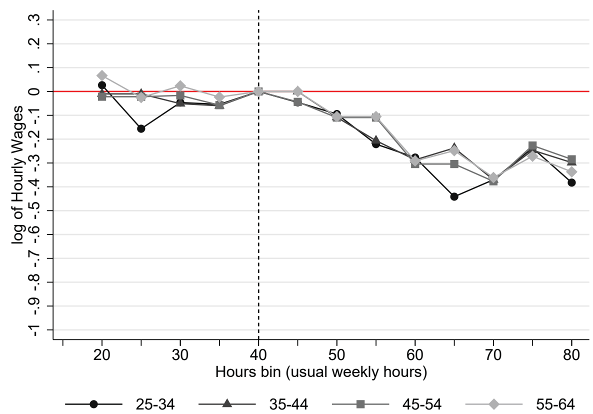
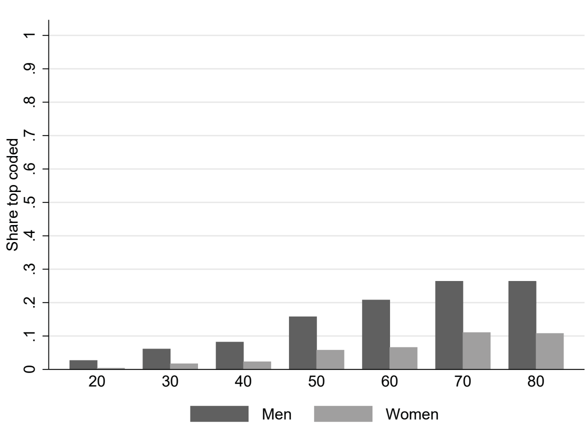


Figure 12: Share with Top-Coded Earnings by Occupation

(a) Lawyers



(b) Physicians and Surgeons

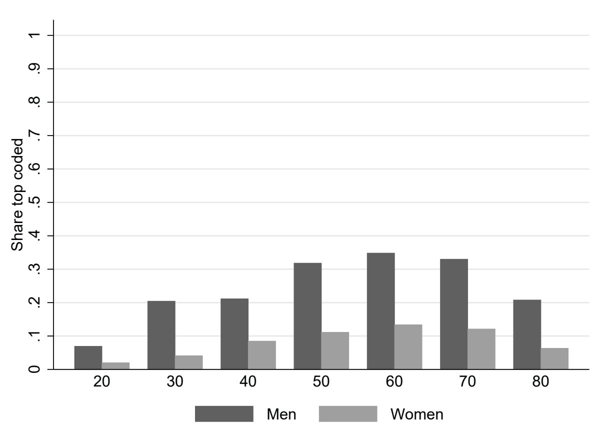
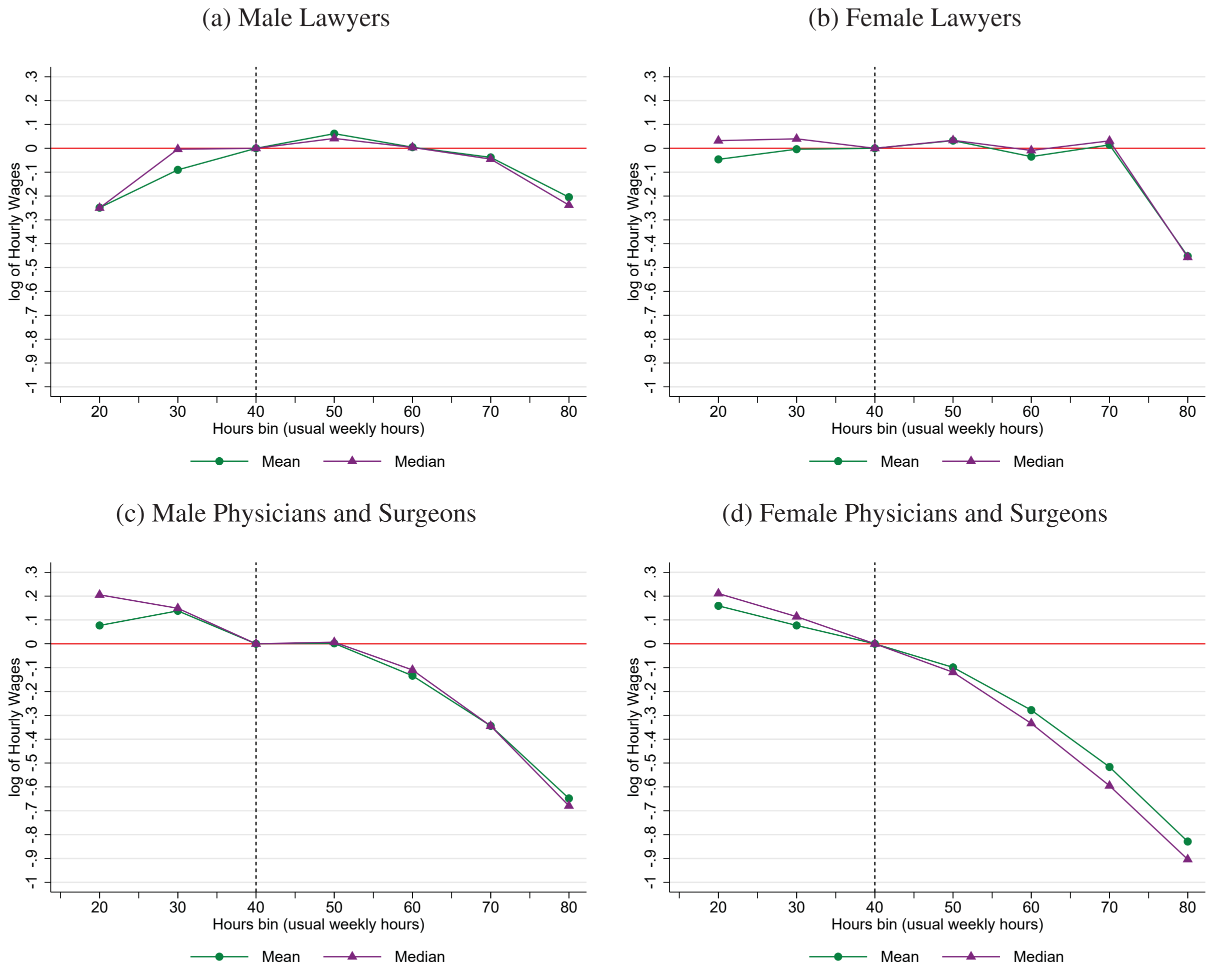


Figure 13: OLS vs. Median Regression Results



hours. The comparison by gender confirm this. While men are much more likely to be top-coded than women, the wage profiles are very similar across both sexes and the OLS-Median regression differences are small for both sexes as well.[11]

To summarize, this section established four key points. First, the wage-hours elasticity varies substantially with hours within occupations, and the correlation between the elasticity in the short hours region and long hours region is fairly small. Second, the elasticity in the short hours region is almost always larger than the elasticity in the long hours region. In particular, for most occupations the wage-hours elasticity in the short hours region is positive, while the wage-hours elasticity in the long hours region is negative. Third, the share of workers working short (long) hours is negatively (positively) correlated with the wage-hours elas-

[11] In the context of a structural model one can implement top-coding in the model consistently with how it is implemented in the data, and thereby implicitly correct for the effects of top-coding. For this correction to be compelling it is important that the model be able to generate the appropriate mass in the right tail of the earnings distribution.

ticity in that region, although quantitatively this correlation is somewhat small, especially for men. Fourth, among the four specific occupations we emphasize, lawyers have a fairly constant wage-hours relationship, while registered nurses, pharmacists, and physicians and surgeons all have fairly constant wages below 40 hours, and wages which decrease with hours above 40 hours.

5 Facts from Panel Analysis

In this section we report some facts based on an analysis of panel data. For this we rely on data from the NLSY79. The NLSY79 allows us to observe workers over nearly their entire career: individuals were first interviewed between ages 14-22, and were 49-56 years old in the most recently available interview wave. Moreover, we can track hours worked over this entire period as these were even asked for during years when individuals were not interviewed.

Replicating our cross-sectional analysis using the NLSY79, we find very similar patterns to those observed in the Census. One noticeable difference, however, is that in the NLSY79 the prevalence of long hours is about five percentage points higher. Results from replicating the cross-sectional analysis using the NLSY are contained in the Appendix.

Although the analysis that we carry out here in this section can be done for both males and females, at this point we only present results from the male subsample.

An important earlier reference from the literature is Imai and Keane (2004). They also used the NLSY79 to estimate a structural model of labor supply, and one of the by-products of their estimation was a structural equation linking hours of work to future wages. Our empirical analysis extends their study along two key dimensions. First, they assumed that the only channel through which hours affected wages was across time, i.e., they did not allow for contemporaneous effects of hours on wages. We show below that abstracting from this effect seems to be quantitatively important.[12]

The second dimension concerns heterogeneity in hours worked. To understand how heterogeneity in hours affects inequality, one must necessarily understand the driving forces behind the dispersion in hours found in the data. The primary focus of Imai and Keane (2004) was to understand how learning by doing

[12]Others have followed a more purely empirical approach. Analogous to the recent literature on cross-sectional relationships, some papers have extended these methods to panel data by regressing current wages on past hours worked, (see, e.g., Gicheva, 2013; and Cortés and Pan, 2017). Similar to the issue we noted in Section 2, a key limitation of this work is that these regressions do not necessarily allow one to infer the causal effect of hours on wages. In addition, they either do not at all control for the effect of current hours or impose a constant elasticity.

would rationalize a high IES even though young workers have low wages and yet do not work that much less than prime age workers. For their purposes the key variation in hours was between younger and older workers, whereas for our analysis the key dimension is overall dispersion.

In what follows we emphasize three patterns from the data. First, we show that controlling for a contemporaneous relationship between hours and wages has a large effect on the estimated dynamic relationship. Second, we present evidence on the persistence of long and short hours over the life cycle. And third, we examine the horizon over which the dynamic effects exert themselves.

5.1 Contemporaneous vs. Dynamic Effects

As just noted, Imai and Keane (2004) assumed that hours worked this period would affect future wages but not wages this period, so that current earnings were linear in current hours worked. The cross-sectional evidence that we presented previously in this proposal suggests not only that this is not the case but also that the contemporaneous effect is quite non-linear across the hours distribution. This suggests that the dynamic estimates from Imai and Keane (2004) might reflect some mix of contemporaneous and dynamic effects. This would potentially have important implications for the joint evolution of hours and wages as a cohort ages. Here we provide some evidence to suggest that this concern is likely to be quantitatively very important.[13]

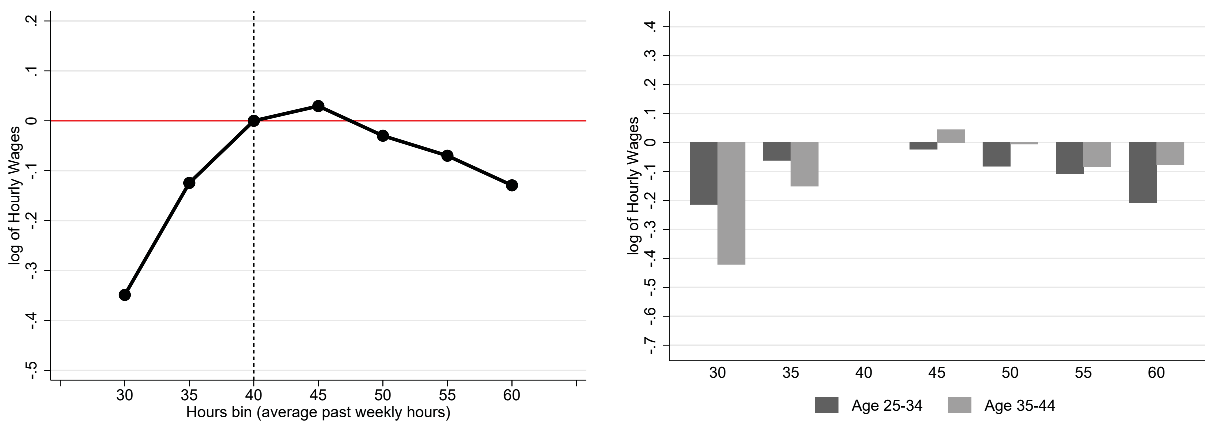
The key mechanism in any learning by doing model is that current wages are a function of past hours of work. We begin by examining this relationship in the NLSY. In particular, for each individual in the NLSY79 at each point in time we compute the average of past usual weekly hours and the current wage rate, defined as total labor earnings divided by total hours. We then run a regression of current log wages on dummies for each of several five hour bins for past usual weekly hours, plus the same controls as before. We then express wages relative to the 40-44 hours bin, as we did previously in the cross-section analysis.

Figure 14 presents the results, both for the entire male sample as well as for two 10-year age groups (all the results in this section hold for 5-year age groups as well). The figure shows that a history of working short hours is associated with substantially lower current wages. We also see that a history of working long hours is associated with lower current wages. For those with previous usual hours in the 50-54 hours bin the effect is relatively modest, but for hours above 60 the negative effect is sizeable. Imai and Keane (2004)

[13] We note one difference between our specification and that in We focus on usual weekly hours whereas they focused on annual hours. For individuals that work full year these two measures are identical, and for males in the NLSY sample the overwhelming majority of them do work full year, so this is likely an issue of second order importance.

Figure 14: log Hourly Wages and Past Hours

(a) All Men (b) By Age



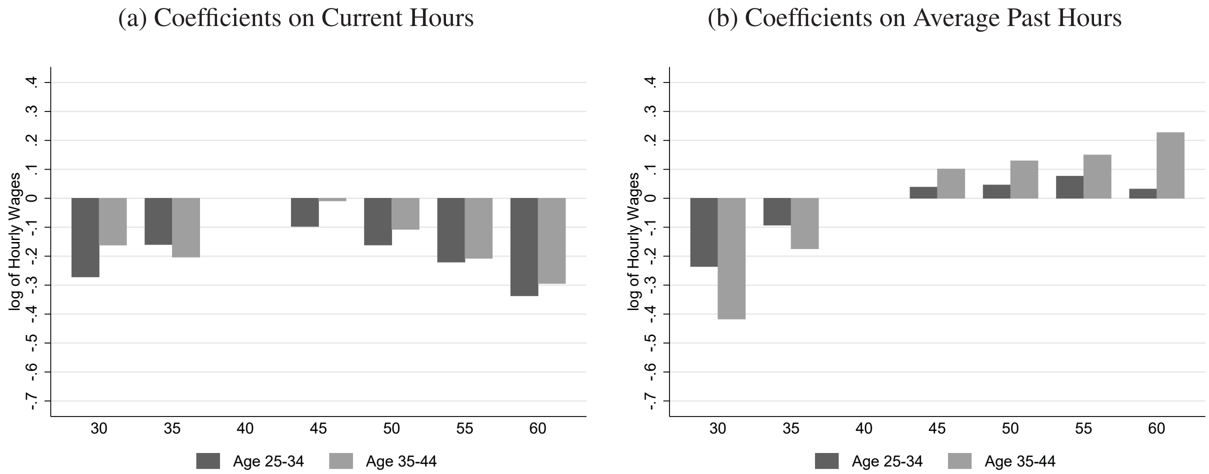
also noted similar features.[14] In particular, their Figure 2 suggested that beyond 2000 hours per year there was no positive effect of additional hours on future wages and that at even higher annual hours the effect turned negative. Based on these features they chose a specification for dynamic returns that mechanically incorporated these patterns.

As for our earlier cross-sectional analysis, we do not interpret the evidence in Figure 14 as reflecting causal dynamic effects. But even absent our earlier evidence on the issue, the age patterns in Figure 14 should also be a cause for caution. For long hours the effects tend to become less negative with age, whereas for short hours the effects become more negative with age. This later fact would be consistent with negative dynamic consequences. But note that there is still a strong negative effect even at very young ages when experience accumulation might be expected to be minimal. It seems quite plausible that the negative effect at young ages might reflect a contemporaneous effect rather than a dynamic effect, since as we show in the next subsection, there is considerable persistence of hours.

To further pursue this, we now generalize the previous regression. In particular, consistent with our cross-sectional analysis we now use log wages as the dependent variable and now also include dummies for current usual weekly hours bins on the right hand side in addition to dummies for past average usual weekly hours. Once again we group observations by age categories. In particular, we run the following regression

[14] Finally, Imai and Keane (2004) adopted a specification in which the return to working longer hours manifested itself in the very next period. This seems somewhat restrictive if one thinks that there are returns to individuals who persistently work long hours and that these returns manifest themselves over time. In our specification we use past average usual hours as our independent variable and we allowed the effect to vary with age. This strikes us as a more flexible way to capture the dynamic effects.

Figure 15: Hourly Wages, Current Hours, and Past Hours



for each of several age bins:

$$w_{it} = \left(\sum_{h \in H} \beta_h \mathbb{1}_{h_{it}} \right) + \beta_{wks} wks_{it} + \gamma X_{it} + \left(\sum_{h \in H} \beta_{exp,\bar{h}} \mathbb{1}_{\bar{h}_{it}} \right) + \varepsilon_{it} \; .$$

Here i denotes an individual in the NLSY79, t denotes the year of the observation between 1979-2014, w denotes log hourly wages, wks denotes log weeks worked, and X is a vector of controls. The $\mathbb{1}_{h_{it}}$ are dummy variables which equal one whenever i's current weekly hours in year t lie in bin h; the $\mathbb{1}_{\bar{h}_{it}}$ are analogous dummy variables for past average weekly hours. Figure 15 shows the estimates for the effect of current and past average hours, again normalizing effects relative to the 40-44 hours bin.

Figure 15a displays the coefficients on current hours. These coefficients display the same qualitative patterns as in our earlier cross-sectional analysis: wages increase as one moves from short to normal hours, and decrease as one moves from normal to long hours. One difference is that wages begin to decrease at a lower hours threshold, at 45-49 hours rather than 50-54 hours.

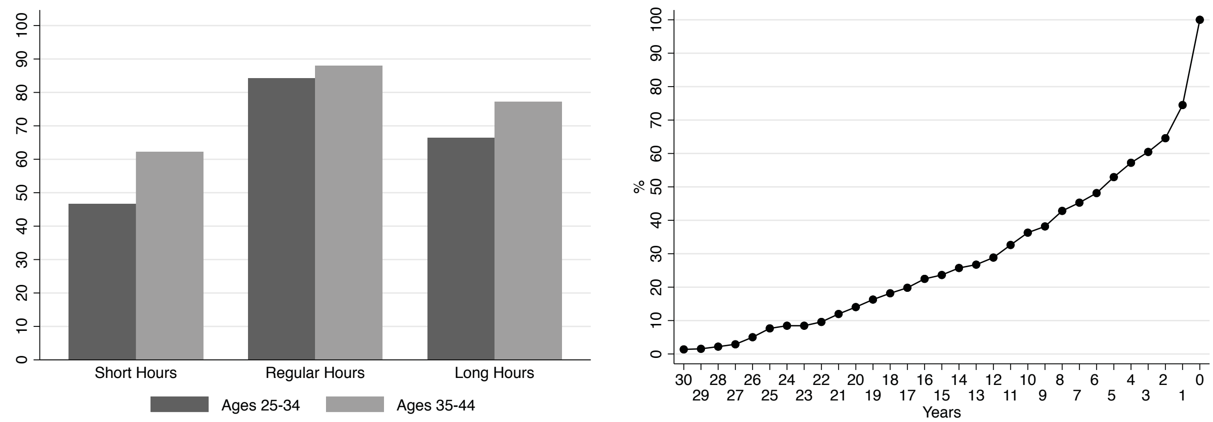
Next we examine the coefficients reflecting the dynamic effects. These are displayed in Figure 15b. The striking feature of this figure is that the effects are negative for short hours but are now positive for long hours. In particular, a history of working long hours is no longer found to be associated with negative effects on current wages, and in fact, for the older age group shown there are very large positive effects associated with working very long hours. More generally, the effect of past hours is larger for the older age group.[15]

[15] These patterns also hold fairly clearly for individuals in the next oldest ten-year age bin, 45-54. However, the size of this group is smaller, and in particular there are very few individuals in the 30-34 and 60-64 hours bins, which increases standard errors and makes it difficult to say anything definitive about these bins.

Figure 16: Hours Persistence

(a) Prob. of Staying in Same Hours Bin between Years t and $t+1$

(b) CDF of Nr. Long Hours Years from a Balanced Panel for Ages 25 to 54



To be sure, the above figures are only reporting some moments from the data and so do not represent causal effects. But our main point is that any attempt to isolate the dynamic effects of hours on wages must also confront the cross-sectional relationship between hours and wages. Failure to do so is very likely to generate large biases in the estimated dynamic effects.

5.2 The Persistence of Long and Short Hours

If there are dynamic effects of hours on wages, then the persistence of heterogeneity in hours is a crucial feature of the data. In particular, if there are high dynamic returns to working long hours, then it is important to capture the extent to which working long hours is a persistent state.

Figure 16a shows the probability that an individual in our NLSY79 sample remains in the same hours region between two subsequent years. For expositional purposes we lump hours into three broad bins: short hours (10-34 hours per week), normal hours (35-49 hours per week), and long hours (50-99 hours per week). For the age group 25-34, about 65% of those working long hours in a given year will also do so the next year and this probability increases to above 75% for the age group 35-44, and even further to 85% for the age group 45-54.

Figure 16b shows how these year to year transitions manifest themselves in accumulated years of long hours experience. We show these results for a balanced panel of men, who we observe over 30 years from age 25 to 54. About 25% of men have worked long hours in at least 15 out of 30 years. At the opposite extreme, about 25% of men experienced zero years of long hours work over 30 years.

The fact that long hours is a persistent state at the individual level suggests that the group of workers who persistently work long hours is likely to be very important in estimating the dynamic return to working long hours. As we alluded to earlier, this is a dimension along which it is important to extend the analysis in Imai and Keane (2004). In their model the only permanent source of preference heterogeneity is captured by the fact that preferences are allowed to vary with educational attainment. But within an educational attainment group, the sole source of preference heterogeneity is shocks that are iid over time. Failure to account for the persistent differences in hours worked within educational attainment categories is likely to interact with estimates of dynamic returns. The reason is that persistent hours differences should have distinctive implications for the evolution of the wage distribution as a cohort ages. Our framework in the next section is static, and therefore there is no distinction between transitory and permanent heterogeneity. However, when we extend our static framework to a dynamic one we will explicitly include permanent heterogeneity with regard to preferences for leisure.

5.3 Time Horizon for Dynamic Effects

At the time of their analysis the oldest individuals in the data set used by Imai and Keane (2004) were only 35 years old. If the goal is to estimate the dynamic returns to working and one thinks that the returns are realized over longer horizons, this might naturally raise concerns. The figures in the previous subsection provide some support for this concern. In particular, when we plotted the estimated effects associated with past hours we noted that the effects were monotonically increasing in past hours for the 35-44 group, but not for the 25-34 group. The data used by Imai and Keane (2004) did not allow them to examine effects over longer horizons.

6 Uncovering the θ_j: A Structural Approach

One of the key facts that we documented earlier in the paper is that there is a systematic and robust relationship between hours and wages in the cross-section. We also presented evidence to suggest that this cross-sectional relationship needs to be understood in order to infer the dynamic effects of hours on wages. As a first step toward isolating the effect of hours on wages, in this section we focus on trying to understand the forces that can generate the cross-sectional patterns documented earlier. Except under very special circumstances, cross-sectional patterns will reflect both dynamic and contemporaneous effects of hours on wages and so do not isolate contemporaneous effects. Nonetheless, in this section we will focus on understanding the extent to which a static model can generate the cross-sectional patterns presented earlier. Despite the fact that results from this exercise would not necessarily carry over directly to a dynamic setting, we adopt this approach as we think it is helpful in allowing us to isolate some of the challenges that arise and is therefore a useful first step.

6.1 Model

We adopt a framework very similar to the one that we used in Section 2. In particular, we assume 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 for work, captured by the parameter α_i, and productivity, which we again represent by z_i. As before, we assume that the two preference parameters σ and γ are the same for all individuals. We now generalize our earlier analysis by assuming that individuals face a schedule for earnings as a function of hours of work that we now represent by:

$$e = zA(h)h^{\theta(h)}$$

We assume that earnings are continuous in h but that $\theta(h)$ is not necessarily continuous. We include the $A(h)$ term as a way to maintain continuity of the earnings function at a point of discontinuity in the $\theta(h)$ function. That is, the function $A(h)$ is constant in any interval in which $\theta(h)$ is continuous, and as a normalization we impose $A(0) = 1$.[16] The appeal of this functional form is that the function $\theta(h)$ provides a clear and flexible mapping from hours into the marginal effect of hours on earnings. We assume no non-labor income and normalize the price of consumption to unity so that each individual has consumption equal to their labor earnings.[17]

It is important to note up-front that even having imposed this amount of structure, one cannot hope to learn anything about the function $\theta(h)$ from cross-sectional data on hours and wages without imposing

[16]We assume that this earnings function is the same for all individuals. In reality there is potential heterogeneity in this function induced by the fact that some but not all workers must be paid overtime for any hours in excess of 40. Using the CPSORG one can distinguish between hourly workers and salaried workers. We can redo our analysis treating these two groups separately. When we do this we obtain very similar results.

[17]We also abstract from taxes at this point, though it would be straightforward to include them.

additional structure. The reason is simple–our model with two dimensions of heterogeneity at the individual level allows us to perfectly account for any cross-sectional pattern of hours and wages even if we assume that $\theta(h)$ is identically equal to one. To see why, note that we could use the wage data to pin down the individual values of the z_i and could use the hours data to pin down the individual values of the α_i. This non-parametric analysis places no restrictions on the joint distribution of the z_i and the α_i. Any attempt to use cross-section data on hours and wages to say something about $\theta(h)$ without direct measures of z must necessarily impose some structure on the nature of heterogeneity.[18]

With this in mind we now impose some structure on the joint distribution of the z_i and the α_i. Specifically, we will assume that they are jointly log normally distributed.[19] This joint distribution is characterized by three values: the mean and standard deviation of $\log z$, the mean and standard deviation of $\log \alpha$, and the correlation between $\log z$ and $\log \alpha$, which we denote by μ_z, σ_z, μ_α, σ_α, and $\rho_{\alpha z}$ respectively. The mean level of z is effectively a normalization, so we normalize μ_z to equal zero.

6.2 Measurement

We will also assume that hours and earnings are both measured with error. Our main concern is that measurement error in hours induces a negative correlation between measured hours and measured wages. Specifically, with classical measurement error, people with high reported hours of work are more likely to have positive measurement error in hours, whereas individuals with low reported hours are more likely to have negative measurement error. This would tend to make wages look artificially low for high hours workers. For this reason we allow for one component of measurement error in log hours that is classical. Another feature of the reported hours distribution is that there is heaping at values ending in either a zero or a five, and in particular there is a particularly large spike at 40 hours. As we noted in Section 2, our framework can generate a spike at 40 hours by assuming that there is a discontinuity, or even a kink in the $\theta(h)$ function at 40. But to capture the fact that people who work "around" 40 hours may be very likely to report that they work exactly 40 hours we will also assume that people with actual hours in the 38-42 range will report

[18]This argument necesssarily generalizes to a dynamic setting as well. If one imposes no structure on the shocks across time then one can trivially account for any panel data on hours and wages without assuming any impact of hours on wages, either contemporaneously or dynamically.

[19]While the assumption of log normal distributions is quite common and convenient in this context, it does imply that we will not generate the fat right tail found in the empirical earnings distribution. We do note, however, that the fat right tail is not as pronounced for labor earnings as it is for overall income.

exactly 40 hours.[20] Note that our distribution for reported hours will still include some individuals who report hours in the interval 38 to 42 without reporting exactly 40 since those with true hours outside of the interval of 38 to 42 are assumed to report with classical measurement error.[21]

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 effect is small. For this reason we abstract from measurement error in earnings in what follows.

6.3 Moment Matching Exercise

We now describe our main quantitative exercise. The idea is to assign values for our model parameters so that the model matches key empirical moments for hours worked and wages. The resulting parametrized model will generate a relationship between hours and wages that includes heterogeneity across workers, measurement error, and the causal effect of hours on wages. We then use our parametrized model to infer the causal impact of hours worked on wages.

A few remarks in order before moving ahead. First, our model implies that all individuals will work positive hours and so there is no selection in terms of who works. We could easily add a selection effect to our model by assuming that each individual has some reservation utility associated with not working; however, we do not think this would change any of our main messages and so have not done this. A richer model would allow the outside option to vary by individual and potentially be correlated with z and α. This generalization would potentially provide an additional reason why productivity and leisure preferences are correlated among workers.

Second, we note that our approach is focused on understanding the patterns in the data from a pure labor supply perspective. That is, we assume that each individual is free to choose their hours of work, but that they may face a tradeoff in terms of hours and wages. We thus abstract from the possibility that an individual who works 40 hours did not have the option to work a different number of hours. Instead, we assume that the wages being offered for other levels of hours were such that the individual preferred to work

[20] As in our empirical analysis we will focus on five hour bins, so that one of our bins will be 40-44. For individuals who actually work either in the interval of 40 to 42 hours this measurement error does not change their bin though it does affect their measured hourly wage.

[21] We can also do our analysis assuming that classical measurement error is the only source of measurement error in hours. As a practical matter this does not have much impact on our findings.

40 hours rather than some other number. To the extent that firms do not find it "profitable" to hire workers for a particular level of hours, we would account for this by having 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. We also note that our earnings function should not be interpreted as the earnings function that a given worker faces at a given firm. Rather, it should be interpreted as the opportunities that the worker faces in the market more broadly. This interpretation is in line with the empirical evidence in Altonji and Paxson (1988). Implicitly, our approach abstracts from any search frictions that a worker might face in finding a job with a particular level of hours.

6.4 How We Assign Parameter Values

In all cases we fix the values of σ and γ. Our exercise can be implemented for any values of these parameters, but in what follows we will focus on the case in which σ tends to one, implying offsetting income and substitution effects, and $\gamma = 0.25$. We discuss later how alternative choices for σ would affect our findings. We note that the exact value of γ is not likely to be important as in our exercise changes in γ will be undone by changes in the standard deviation of the preference shocks. In what follows we also fix the level of classical measurement error at $\sigma_m = 0.08$.[22]

For the results reported here we impose a functional form for $\theta(h)$. Our choice is informed by the features of the evidence reported earlier. Specifically we assume that $\theta(h)$ is constant below 40 hours and that it is linear after 40 hours. We allow for a discontinuity at 40, which can generate a spike in the share of workers working 40 hours. A constant elasticity provides a good fit to the wage-hours profile between 30-40 hours, and there relatively few workers below 30 hours.[23] Above 40 hours, rather than imposing a constant θ we allow $\theta(h)$ to vary linearly in order to simultaneously generate two patterns: (i) a gradual decline in the slope of the wage-hours profile above 40 hours, and (ii) no spike in the share of workers in any bin above 40 hours.

While this specification imposes quite a bit of structure we will see that it is sufficiently flexible to account for the features of the data that we target. This choice implies that $\theta(h)$ is characterized by three

[22]There is no definitive value for the extent of measurement error in hours. Assuming that measurement error is classical and iid over time then transitory variation in hours provides some information about plausible values. We adopt this value based on the estimates in Heathcote et al. (2014) regarding the variance of the transitory component of hours. Duncan and Hill (1985) and Bound et al. (1994) are two examples of small scale studies documenting discrepancy between adminsitrative data and survey responses. They find even larger estimates of measurement error in hours. See also the survey article by Bound et al. (2001).

[23]It would be relatively straightforward to modify the $\theta(h)$ function below 30 hours so as to provide a better fit to the wage data in that region. We postpone doing so here simply to avoid the additional parameters.

Table 4: Targets for Overall Sample

Moment	Males	Females
$sd\ \log w$	0.70	0.67
$sd\ \log h$	0.23	0.30
$\exp(mn \log h)$	43.8	38.1

parameters: the constant value below 40 hours, denoted by θ_1, and the intercept and slope term for the linear portion, denoted by θ_2 and θ_3 respectively. As noted earlier, any discontinuity in the $\theta(h)$ function at 40 is offset by a change in $A(h)$ that makes earnings continuous. The restrictions on $A(h)$ and the requirement that earnings be continuous implies that $A(h)$ is completely determined by the $\theta(h)$ function.

Our choices up to this point leave seven parameters whose values are not yet assigned: the four parameters for the joint distribution of z and α (recall that we normalized μ_z to equal 0), and the three parameters for the earnings function. As a first step we seek values of these parameters that can reproduce the following features of the data: the standard deviation of log wages, the standard deviation of log hours, the mean of log hours, and the profile of mean wages by five hour bins. At this stage we do not use any additional moments of the hours distribution, but in what follows we will examine the extent to which we replicate the distribution of workers across 5-hour bins. We view this as a simple and transparent way to connect the model with the data and plan to explicitly incorporate additional moments of the hours distribution into future iterations.

Table 4 reports the standard deviation of log hourly wages, the standard deviation of log hours, and the mean of log hours in our overall male and female samples. [24] We implement the exercise above with the following procedure:

1. Fix the correlation parameter $\rho_{\alpha,z}$, and define a three-dimensional grid over the earnings function parameters $(\theta_1, \theta_2, \theta_3)$.

2. Given values for $\rho_{\alpha z}$ and $(\theta_1, \theta_2, \theta_3)$, choose values for the remaining distribution parameters $(\sigma_z, \sigma_\alpha, \mu_\alpha)$ to exactly match the standard deviation of log wages, the standard deviation of log hours, and the mean of log hours. Intuitively, all else equal, (i) increasing σ_z increases the standard deviation of log wages; (ii) increasing σ_α increases the standard deviation of log hours; and (iii) increasing μ_α decreases the mean of log hours.

[24] Specifically, we analyze mean wages within the following 5-hour bins: 30-34, 35-39, ..., 60-64. We focus on mean wages for these hours because roughly 90% of our overall male and female samples fall into this range.

Table 5: Parameter Values To Match Overall Male and Females Samples, Given $\rho_{\alpha z} = 0$

Parameter Name	Males	Females
σ_z	0.68	0.65
μ_α	-18.8	-18.11
cv_α	0.08	0.11
θ_1	2.05	1.72
θ_2	2.51	3.30
θ_3	-0.0078	-0.0120

3. Given $\rho_{\alpha,z}$, choose the combination of $(\theta_1, \theta_2, \theta_3)$ that minimizes the sum of squared residuals between the profile of mean wages by five hour bins in the model and data.

Given a value of $\rho_{\alpha z}$, the remaining model parameters are uniquely identified by our targeted moments. However, as we document in the next section, these moments are insufficient to pin down a single value of $\rho_{\alpha z}$. One might think that the correlation between wages and hours would provide information about this parameter. However, we find that varying $\rho_{\alpha z}$ has essentially no effect on the model's wage-hours correlation conditional on the model matching the rest of the targeted moments. This is because our model already fits the hours distribution fairly well, and also matches the standard deviations of wages and the profile of mean wages by hours bins, which essentially pins down the correlation between wages and hours.

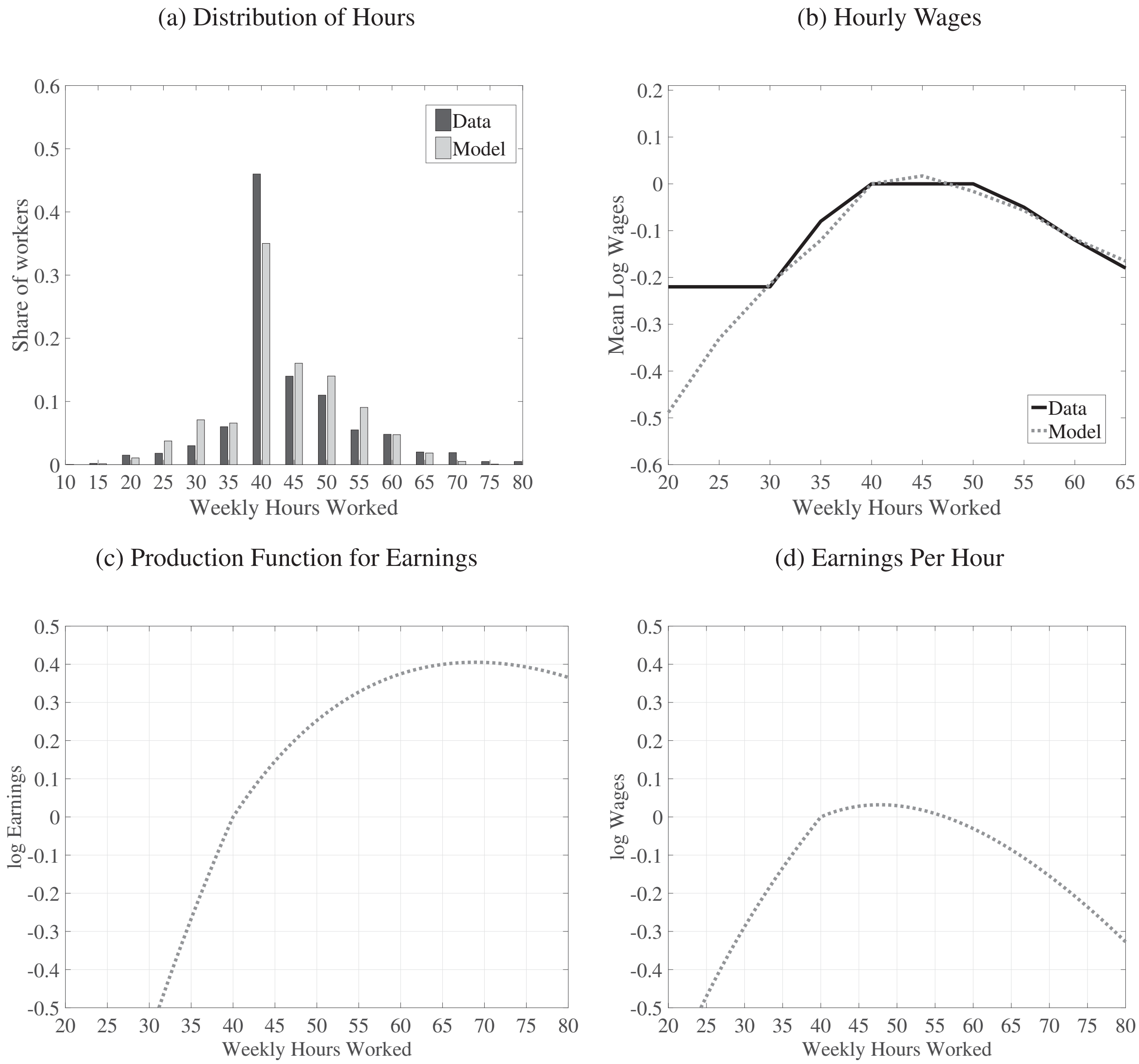
For this reason, we exogenously impose multiple values for $\rho_{\alpha z}$ and document the range of the rest of the implied parameter values. At the end of Section 6.5 we outline one strategy for using external empirical evidence to allow us to significantly narrow the range of plausible values for $\rho_{\alpha z}$.

The above exercise can be implemented on various cuts of the data. We first report results based on the overall sample by gender. We then consider four specific occupations for men: lawyers, pharmacists, physicians and surgeons, and registered nurses.

6.5 Results for the Overall Population of Men

Overall the model generates a tight fit with the cross-sectional features of the data for our male sample. To illustrate this, we first present the case where leisure preferences α and productivity z are uncorrelated, i.e., $\rho_{\alpha z} = 0$. The resulting calibrated parameter values are listed in Table 5. Figure 17a shows the fit of the model to the hours worked distribution. The model matches the distribution of hours fairly well. In particular, the model generates a large spike in the share working 40-44 hours, although the spike is still somewhat smaller than in the data. Two model features produce this spike. First, the earnings function is

Figure 17: Fit of Calibrated Model for $\rho_{\alpha z} = 0$: Male Hours and Hourly Wages



Note: Hours worked in the model correspond to reported hours (they include measurement error). Hourly wages in the model are computed by dividing earnings by reported hours.

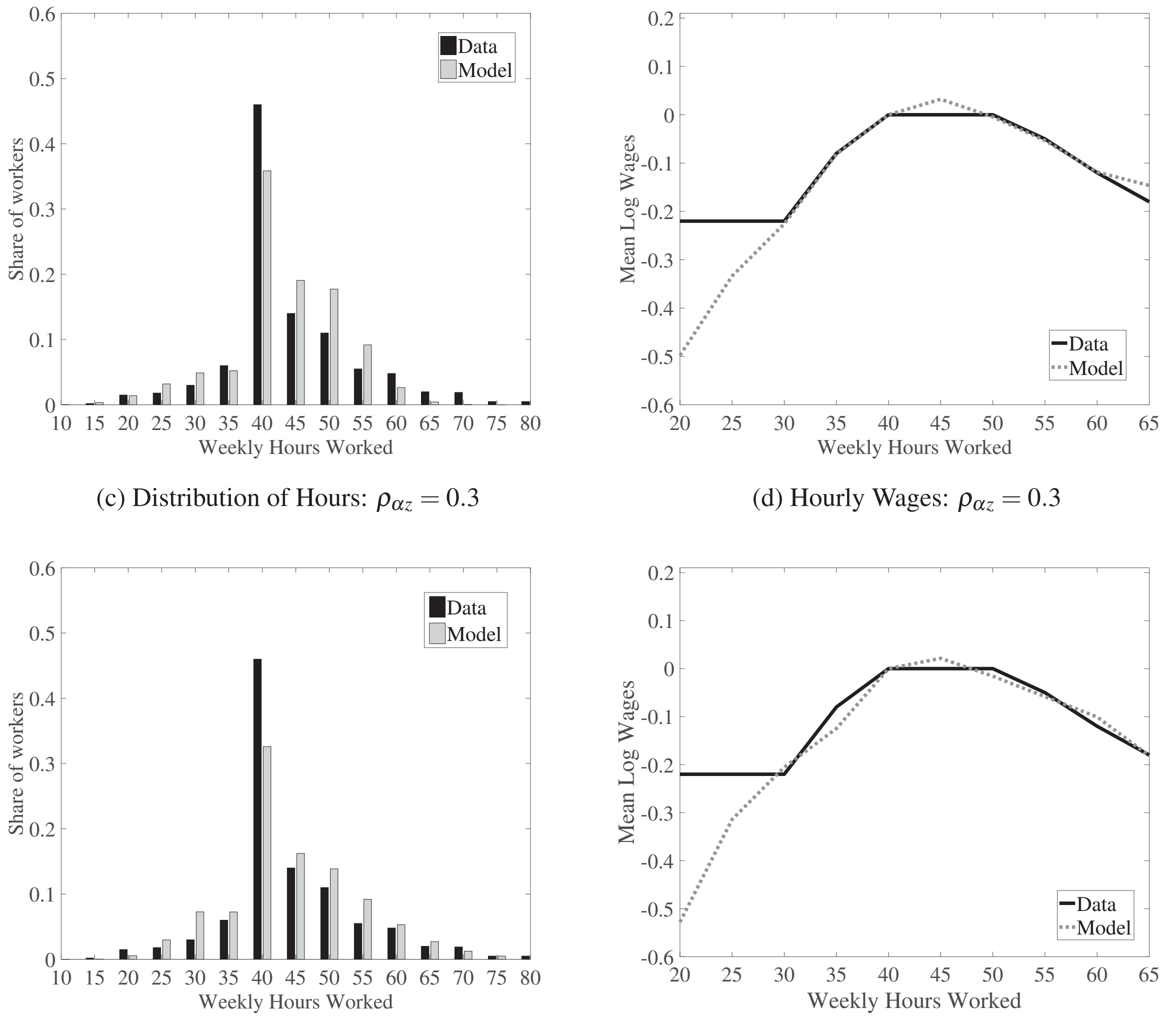
kinked at 40 hours (see Figure17c), which leads to bunching of workers. Second, conditional on the actual distribution of workers, our specification of measurement error increases the share of workers who report working 40 hours. Quantitatively, we find that the kink at 40 plays the largest role in generating the spike.

Figure 17b shows the fit of the model to the wage-hours profile. The model fits the data profile tightly between 30 and 69 hours per week, which is the range we target. The model produces a positive wage-hours slope below 40 hours because $\theta_1 > 0$. From 40-50 hours the model's wage profile is fairly flat, but beyond 50 hours the slope is negative. The model generates this pattern because $\theta_2 > 1$, but $\theta_3 < 0$, which implies $\theta(h)$ declines as hours increase beyond 40. Figure 17c traces out the earnings function implied by these

Figure 18: Fit of Calibrated Model for $\rho_{\alpha z} = -0.1$ and 0.3

(a) Distribution of Hours: $\rho_{\alpha z} = -0.1$ (b) Hourly Wages: $\rho_{\alpha z} = -0.1$

(c) Distribution of Hours: $\rho_{\alpha z} = 0.3$ (d) Hourly Wages: $\rho_{\alpha z} = 0.3$



Note: Hours worked in the model correspond to reported hours (they include measurement error). Hourly wages in the model are computed by dividing earnings by reported hours.

parameter values, and Figure 17d shows the corresponding earnings per hour.[25]

In the above scenario, with $\rho_{\alpha z} = 0$, the reported hourly wage profile in Figure 17b and the actual hourly wage profile in Figure 17d differ from one another only because of measurement error in hours. However, in general these two profiles may also differ from one another due to heterogeneous worker productivity that is correlated with hours.

To illustrate this we repeat our moment-matching exercise for different values of $\rho_{\alpha z}$. On the upper end we examine a fairly high correlation of $\rho_{\alpha z} = 0.3$. This implies that more productive workers tend

[25]The model wage profile fits the data poorly below 30 hours. We could easily improve this by allowing the elasticity to change at 30 hours as well as at 40 hours. However, given the small share of workers below 30 hours, we postpone this step for now.

Table 6: Male Hourly Wages Relative to $h = 40$

	$h = 30$	$h = 40$	$h = 50$	$h = 60$
$\rho_{\alpha z} = -0.1$	-0.19	0.00	-0.07	-0.21
$\rho_{\alpha z} = 0.0$	-0.29	0.00	0.03	-0.03
$\rho_{\alpha z} = 0.1$	-0.40	0.00	0.11	0.11
$\rho_{\alpha z} = 0.2$	-0.51	0.00	0.21	0.30
$\rho_{\alpha z} = 0.3$	-0.63	0.00	0.32	0.52

to have high disutility from work and therefore work fewer hours. On the lower end we only go as low as $\rho_{\alpha z} = -0.10$. With a negative correlation, more productive workers tend to work longer hours, so to generate low wages at long hours the calibration procedure must assign a very low value of θ in this region. As $\rho_{\alpha z}$ becomes sufficiently negative, the required θ becomes *negative*, which says that working long hours *lowers* total earnings. In this case no workers choose to work long hours, which effectively puts a lower bound on $\rho_{\alpha z}$.

Figure 18 displays the model fit for hours and wages given $\rho_{\alpha z} = -0.1$ and $\rho_{\alpha z} = 0.3$. Combined with Figure 17, these figures demonstrate that the model produces a similarly good fit for a wide range of values of $\rho_{\alpha z}$. Thus, one key result from this exercise is that cross-sectional data on the joint distribution of hours and wages is not sufficient to uniquely identify the correlation between unobserved heterogeneity and hours.

We now show that a corollary of the above result is that cross-sectional data on hours and wages is insufficient to identify the shape of $\theta(h)$, and therefore is insufficient to identify the wage impact of working short or long hours. Table 6 displays the implied wage penalties and premia associated with working short and long hours for several values of $\rho_{\alpha z}$. Relative to working 40 hours, log hourly wages at 30 hours range from -0.19 to -0.71. On the other end, relative log hourly wages at 60 hours range from -0.25 to 0.65. There is an intuitive pattern that appears in Table 6.

One way to see this is to consider the wage-hours profile that would result if we set $\theta(h) = 1$ for all h and allowed the value of $\rho_{\alpha z}$ to vary. If $\rho_{\alpha z}$ is negative then people with high productivity tend to work long hours, so the wage-hours profile would be upward sloping. If $\rho_{\alpha z}$ is positive, then people with high productivity tend to work shorter hours and we would see a downward sloping profile. And if $\rho_{\alpha z} = 0$ then the wage-hours profile will be flat. Given that the wage-hours profiles measure wages relative to those in the 40-44 hours bin, it follows that changing the value of $\rho_{\alpha z}$ essentially rotates the wage-hours profile with the pivot point being the 40-44 hours bin.

Now, start with a specification that matches the data assuming that $\rho_{\alpha z} = 0$. Loosely speaking, as we consider alternative values of $\rho_{\alpha z}$ the direct effect would be to pivot the wage-hours profile. In order to continue to match the data, the $\theta(h)$ function would have to adjust to basically offset the pivot induced by the new value of $\rho_{\alpha z}$. In fact, what we see in the above table is exactly this form of pivoting in the $\theta(h)$ function. If $\rho_{\alpha z}$ is negative, thereby inducing a negative slope in the hours-wage profile, the $\theta(h)$ function would need to induce a higher return for high hours and a lower return for lower hours in order to offset this effect. In other words, generating a more positive (or less negative) effect at higher hours requires generating a larger negative (or smaller positive) effect at low hours. Put somewhat differently, as we vary $\rho_{\alpha z}$ we basically pivot the implied effects of $\theta(h)$ around the $h = 40-44$ bin.

We conclude this section with a potentially useful implication of this relationship between $\rho_{\alpha z}$ and $\theta(h)$. Our moment matching exercise for an interval of values for $\rho_{\alpha z}$ basically determines a family of $\theta(h)$ curves that differ in terms of their rotation around a particular hours bin (with our normalization, this is the 40-44 hours bin). It follows that if we know the value of "true" wages, i.e. the one implied by the actual earnings function, relative to the 40-44 hours bin at just one other level of hours we could identify the entire curve. In this regard the evidence in Aaronson and French (2004) is potentially valuable. They present quasi-experimental evidence that provides an estimate of the part time wage penalty for older males. They find a wage penalty on the order of 25%. Their analysis did no provide any information about the potential wage impact of working longer hours. But based on our previous discussion, when this one piece of evidence on the wage impact of short hours is combined with our structural analysis, we can effectively pin down the entire profile for θ. Focusing on the range of values for $\rho_{\alpha z}$ that generate a part time wage penalty on the order of 25% would allow us to limit attention to roughly the first two rows in Table 6. As the table shows, doing so narrows the range of potential wage premia at 50 hours to $[-0.11, 0.03]$, and at 60 hours to $[-0.25, -0.03]$. Interestingly, this approach rules out large wage gains for working long hours.[26]

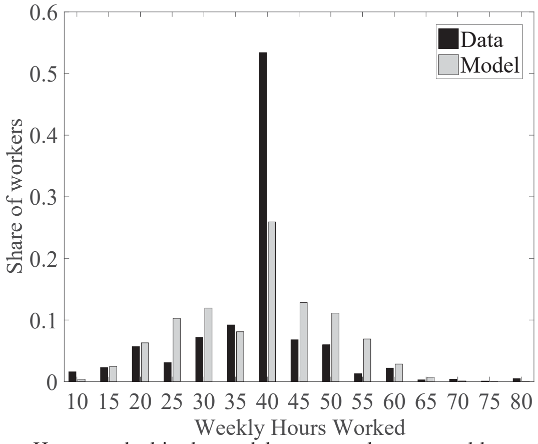
6.6 Results for Overall Female Sample

Our model can also generate a good fit to the cross-sectional features of the overall sample of females. Figure 19 displays the distribution of hours and the wage-hours profile in the model and data. Relative to men, for females hours are shifted left and the standard deviation of log hours is larger. One empirical

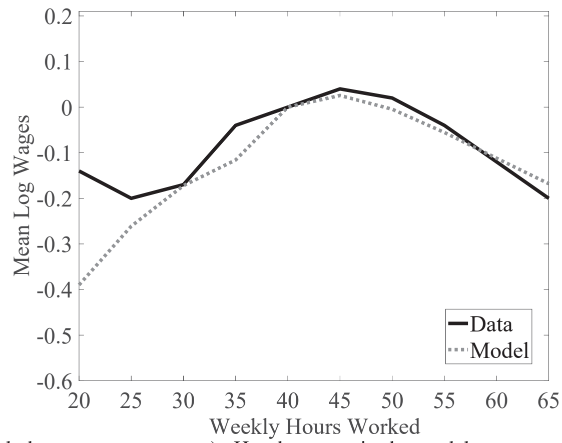
[26] Although we do not pursue it here, the relationship between $\rho_{\alpha z}$ and $\theta(h)$ also suggests how assuming different values for σ would affect the results. As we deviate from $\sigma = 1$ we no longer have offsetting income and substitution effects, and this would induce a slope to the wage-hours profile even if $\theta(h) = 1$ for all h. Intuitively, this induces effects that are very similar to changes in $\rho_{\alpha z}$ and we expect that the results would mirror those just presented.

Figure 19: Fit of Calibrated Model for $\rho_{\alpha z} = 0$: Female Hours and Hourly Wages

(a) Distribution of Hours



(b) Hourly Wages



Note: Hours worked in the model correspond to reported hours (they include measurement error). Hourly wages in the model are computed by dividing earnings by reported hours.

pattern the model struggles to match is that the spike in the hours share at 40 is larger for females than for males, while in the model this is reversed. The model wage-hours profile fits the empirical profile fairly tightly in the 30-69 hour range that we target, although model wages are a bit too low in the 35 hour bin.

As with our sample of men, the moment matching exercise is unable to uniquely identify the value of $\rho_{\alpha z}$, and is therefore unable to uniquely identify the function $\theta(h)$. Table 7 displays the wage penalties and premiums for females associated with working short and long hours, assuming different values of $\rho_{\alpha z}$. Qualitatively we observe the same intuitive pattern for women as we did for men: larger values of $\rho_{\alpha z}$ monotonically increase the implied long hours premium and exacerbate the short hours penalty. Quantitatively, relative to Table 6 for men the range of implied short hours penalties and long hours premiums is narrower. For example, allowing $\rho_{\alpha z}$ to vary from -0.1 to 0.3 changes the wage impact of working 30 hours from -0.19 to 0.63 for men, but from -0.14 to -0.35 for women.

While we view these aggregate analyses for men and women as an informative first step, it is well known that there are important interactions between occuapations, hours, wages. In particular, Goldin (2014) argues that one important reason that gender wage gaps are higher in some occupations than other is because the mapping between hours and wages differs across occupations. To begin to explore the implications of our structural model for these claims, in the next section we repeat our moment matching exercises for men in different 3-digit occupations.

Table 7: Female Hourly Wages Relative to $h = 40$

	$h = 30$	$h = 40$	$h = 50$	$h = 60$
$\rho_{\alpha z} = -0.1$	-0.14	0.00	-0.01	-0.13
$\rho_{\alpha z} = 0.0$	-0.20	0.00	0.04	-0.05
$\rho_{\alpha z} = 0.1$	-0.24	0.00	0.11	0.10
$\rho_{\alpha z} = 0.2$	-0.31	0.00	0.15	0.15
$\rho_{\alpha z} = 0.3$	-0.35	0.00	0.24	0.27

6.7 Results for Selected Occupations

In this subsection we apply our method to data for males in the four occupations that we had considered earlier in the paper: lawyers, pharmacists, registered nurses, and surgeons. We note up-front that imposing a functional form on the joint distribution of productivity and preferences for work is perhaps much more delicate in this context, since we might expect that selection effects are such that each occupation draws from a particular slice of the overall distribution, making it unlikely that the distribution within an occupation might be of the same general form as the overall distribution. But having offered this qualification, we proceed with our assumption of joint log normality. Going forward one could embed our analysis into a model of occupational choice, so that the selection effects are determined as part of the analysis.

Figures 20–23 display the fit of the model to the hours distributions and wage-hours profiles of each occupation. We find that our model is able to achieve a reasonable fit to the hours distribution and wage-hours profiles for all four occupations. This is particularly encouraging because the empirical targets for hours and wages differ substantially across occupations. For example, the share working 50 hours or more varies from 0.174 for registered nurses to 0.766 for physicians and surgeons. To take another example, wages in the 50 hours bin are 6.1 log points higher than in the 40 hours bin for lawyers, but for pharmacists relative wages are 11.5 log points lower.

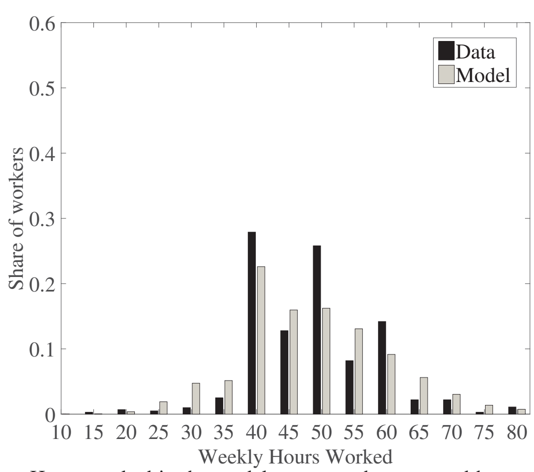
As with the overall samples of men and women, our moment matching exercise presents us with a family of estimated $\theta(h)$ functions as we vary the correlation of individual characteristics. Also similar to our overall samples of men and women, Tables 8 and 9 document that the implied short hours and long hours penalties/premiums vary substantially as we change $\rho_{\alpha z}$. As noted earlier in the exercise focused on aggregate patterns we think that information on higher moments of the hours distribution may be help to distinguish between these different specifications. In particular, we note that the difference between the long and short hours shares do vary across these specifications and so might be a valuable dimension.

Our eventual goal is to identify an earnings function $\theta(h)$ on an occupation-by-occupation basis. This would allow us to identify the joint effect of occupations and hours on wages and wage inequality, both within and between genders. One potential path toward this end would be to replicate the spirit of the procedure in Aaronson and French (2004) on an occupation-by-occupation basis. That is, to use quasi-experimental variation in hours to identify wages at one point in the hours distribution relative to the 40-44 hour bin, and then to choose the corresponding $\rho_{\alpha z}$ and $\theta(h)$.

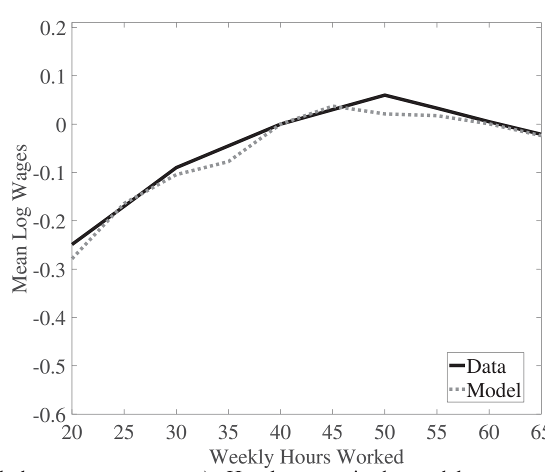
For now we note that the findings just presented are potentially consistent with the narrative offered by Goldin (2014). Specifically, suppose one had the prior that the wage penalty for short hours is relatively small for nurses and substantial for surgeons. Given this prior on the wage impact of working *short hours*, the results in Tables 8 and 9 predict that working *long hours* would have a small or negative impact on wages for nurses, but a substantial positive impact on wages for surgeons. Our framework also makes predictions about some additional properties that are part of supporting these priors. In particular, via Tables 8 and 9 a minor short hours penalty for nurses implies a small or positive correlation between unobserved productivity and hours for that occupation, while a sizable part time penalty for surgeons implies a strong negative correlation between productivity and hours in that occupation. Of course, while our findings are consistent with this narrative, they cannot be viewed as definitive support for this narrative absent additional information that would allow us to single out a particular value for the correlation of individual characteristics.

Figure 20: Fit of Calibrated Model for Male Lawyers, with $\rho_{\alpha z} = 0$: Hours and Hourly Wages

(a) Distribution of Hours



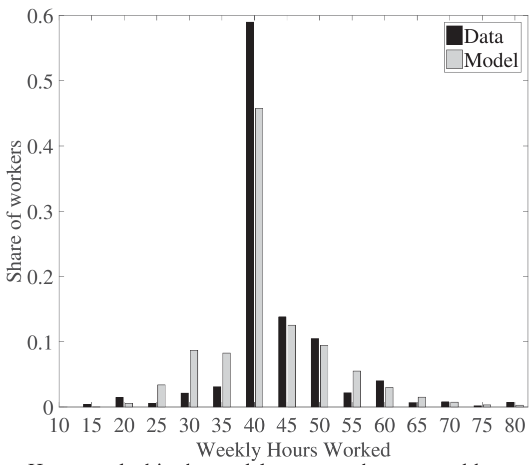
(b) Hourly Wages



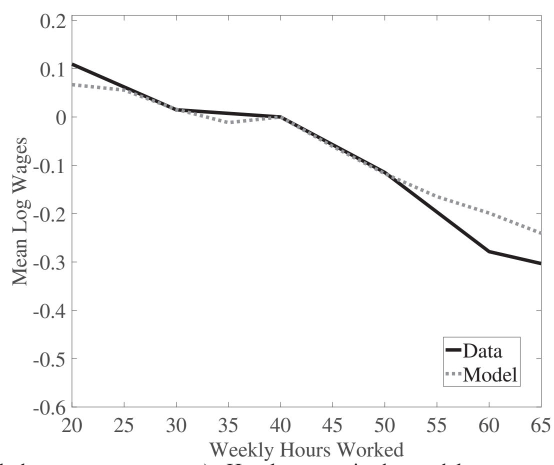
Note: Hours worked in the model correspond to reported hours (they include measurement error). Hourly wages in the model are computed by dividing earnings by reported hours.

Figure 21: Fit of Calibrated Model for Male Pharmacists, with $\rho_{\alpha z} = 0$: Hours and Hourly Wages

(a) Distribution of Hours



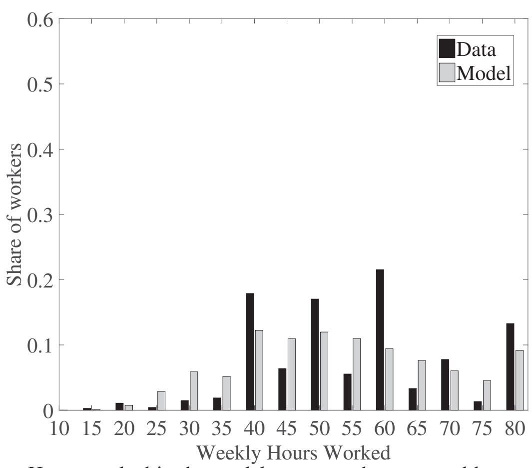
(b) Hourly Wages



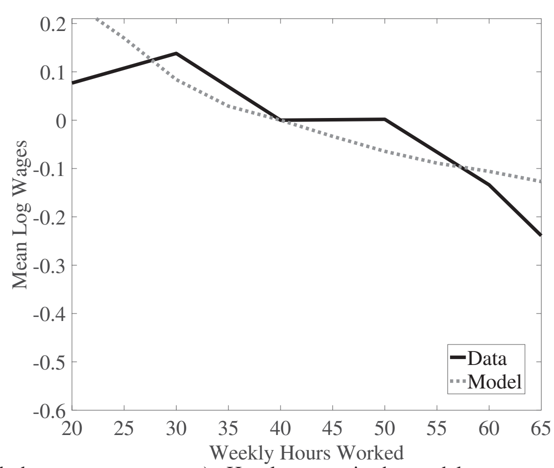
Note: Hours worked in the model correspond to reported hours (they include measurement error). Hourly wages in the model are computed by dividing earnings by reported hours.

Figure 22: Fit of Calibrated Model for Male Physicians and Surgeons, with $\rho_{\alpha z} = 0$: Hours and Hourly Wages

(a) Distribution of Hours



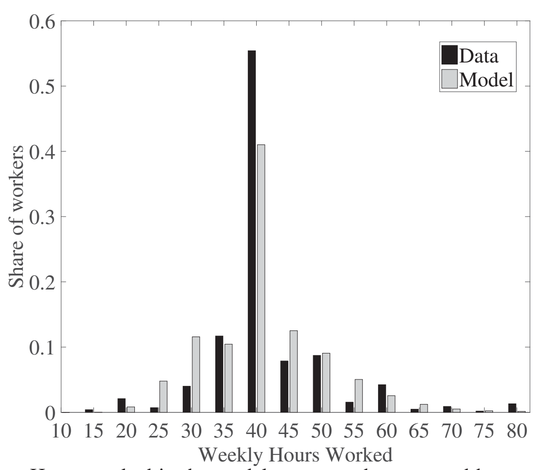
(b) Hourly Wages



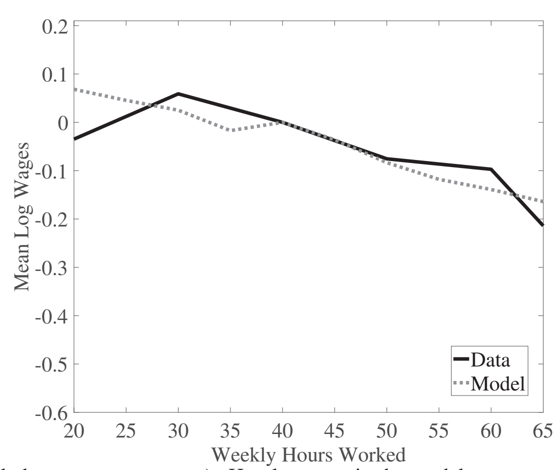
Note: Hours worked in the model correspond to reported hours (they include measurement error). Hourly wages in the model are computed by dividing earnings by reported hours.

Figure 23: Fit of Calibrated Model for Male Registered Nurses, with $\rho_{\alpha z} = 0$: Hours and Hourly Wages

(a) Distribution of Hours



(b) Hourly Wages



Note: Hours worked in the model correspond to reported hours (they include measurement error). Hourly wages in the model are computed by dividing earnings by reported hours.

Table 8: Hourly Wages Relative to $h = 40$: Physicians

	$h = 30$	$h = 40$	$h = 50$	$h = 60$
$\rho_{\alpha z} = -0.1$	0.16	0.00	-0.07	-0.17
$\rho_{\alpha z} = 0.0$	0.08	0.00	-0.04	-0.08
$\rho_{\alpha z} = 0.1$	0.03	0.00	0.01	0.01
$\rho_{\alpha z} = 0.2$	0.00	0.00	0.05	0.08
$\rho_{\alpha z} = 0.3$	-0.03	0.00	0.09	0.16

Table 9: Hourly Wages Relative to $h = 40$: Registered Nurses

	$h = 30$	$h = 40$	$h = 50$	$h = 60$
$\rho_{\alpha z} = -0.1$	0.06	0.00	-0.09	-0.16
$\rho_{\alpha z} = 0.0$	0.00	0.00	-0.05	-0.09
$\rho_{\alpha z} = 0.1$	-0.05	0.00	0.05	0.08
$\rho_{\alpha z} = 0.2$	-0.10	0.00	0.08	0.14
$\rho_{\alpha z} = 0.3$	-0.14	0.00	0.12	0.20

7 Conclusion

Our goal in this project is to learn more about the effects of hours of work on hourly wages, and how this effect differs across occupations. Our initial contribution is to document several novel facts about the relationship between hours and wages using both cross-sectional and panel data. Relative to the existing literature we offer two key findings.

First, we show that the cross-sectional relationship between hours and wages is not well captured by a single elasticity. That is, the elasticity of mean (and median) hourly wages with respect to hours varies with the level of hours worked. In our aggregate sample wages increase with hours below 40 hours, are roughly constant in hours between 40-50 hours, and decline sharply with hours beyond 50 hours. This pattern remains robust when we partition our sample by gender, education, and age. By contrast, when we partition our sample by occupation we document substantial heterogeneity in the cross-sectional relationship between hours and wages. However, qualitatively consistent with our other results, we find that in most occupations wages increase with hours below 40 hours and decrease with hours above 50 hours. Interestingly, the correlation between an occupation's wage-hours elasticity below 40 hours and its elasticity above 50 hours is small, implying large variation in the "shape" of the wage-hours relationship across occupations.

Second, we show that the contemporaneous relationship between hours and wages looks very different from the dynamic relationship. In particular, when we control for both current hours and average past hours, long current hours continure to be associated with lower wages, but long past hours are associated with higher wages. One important implication of this finding is that any attempt to analyze the dynamic effects of hours on future wages must also *simultaneously* address the contemporaneous relationship. The reason is that hours are persistent over time for a given individual, implying that current hours are positively correlated with past hours. Therefore, comparing wages to past hours while neglecting to control for current hours introduces a bias in the coefficient of past hours. More generally, our dynamic findings suggest that future benefits may be an important determinant of the hours decision, and that a thorough understanding of the effect of hours on wages must address this dynamic component.

We then explore the extent to which these empirical findings allow one to make inference about the causal effect of hours on wages. The first point we emphasize is that one cannot make any such inference without additional structure. The reason is that one cannot disentangle the causal effect of hours on wages from the role of unobserved heterogeneity correlated with hours. This finding motivates us to present a structure and method that allows us to make some inference about these causal effects. But even with the considerable structure that we impose, at this point our method is only able to isolate a family of potential effects rather than a single effect. We suggest one source of outside information that might be useful for providing additional identification, and plan to pursue this in future work.

References

AARONSON, D. AND E. FRENCH (2004): "Do Part-time Workers Have Lower Wages than Full-Time Workers? Evidence from Social Security Budget Constraints," *Journal of Labor Economics*, 22, 329–352.

ACEMOGLU, D. AND D. AUTOR (2011): *Handbook of Labor Economics*, Elsevier, vol. 4, chap. Skills, Tasks and Technologies: Implications for Employment and Earnings, 1043–1171.

ALTONJI, J. G. AND C. H. PAXSON (1988): "Labor supply preferences, hours constraints, and hours-wage trade-offs," *Journal of Labor Economics*, 6, 254–276.

BERTRAND, M. (2017): "The Glass Ceiling," *Economica*, forthcoming.

BERTRAND, M., C. GOLDIN, AND L. F. KATZ (2010): "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, 2, 228–255.

BOUND, J., C. BROWN, G. J. DUNCAN, AND W. L. RODGERS (1994): "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, 12, 345–368.

BOUND, J., C. BROWN, AND N. MATHIOWETZ (2001): *Handbook of Econometrics*, Elsevier, vol. 5, chap. Measurement Error in Survey Data, 3705–3843.

CORTES, P. AND J. PAN (2016a): "Prevalence of Long Hours and Skilled Women's Occupational Choices," .

——— (2016b): "When Time Binds: Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled," *Journal of Labor Economics*, forthcoming.

CORTÉS, P. AND J. PAN (2017): "When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap," *Journal of Labor Economics*, forthcoming.

DUNCAN, G. J. AND D. H. HILL (1985): "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data," *Journal of Labor Economics*, 3.

EROSA, A., L. FUSTER, G. KAMBOUROV, AND R. ROGERSON (2017): "Hours, Occupations, and Gender Differences in Labor Market Outcomes," Tech. Rep. 23636.

GICHEVA, D. (2013): "Working Long Hours and Early Career Outcomes in the High-End Labor Market," *Journal of Labor Economics*, 31, 785–824.

GOLDIN, C. (2014): "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, 104, 1091–1119.

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. J. (1976): *Household production and consumption*, Columbia University Press, New York, chap. Estimates of a human capital production function embedded in a life-cycle model of labor supply, 227–264.

IMAI, S. AND M. P. KEANE (2004): "Intertemporal labor supply and human capital accumulation*," *International Economic Review*, 45, 601–641.

MOFFITT, R. (1984): "The estimation of a joint wage-hours labor supply model," *Journal of Labor Economics*, 2, 550–566.

ROSEN, H. (1976): "Taxes in a Labor Supply Model with Joint Wage-Hours Determination ," *Econometrica*, 44, 485–507.

SHAW, K. L. (1989): "Life-cycle labor supply with human capital accumulation," *International Economic Review*, 431–456.

A Cross-Sectional Results for the NLSY

Figure 24: Fraction Working Short, Long Hours and Very Long Hours

(a) Men (b) Women

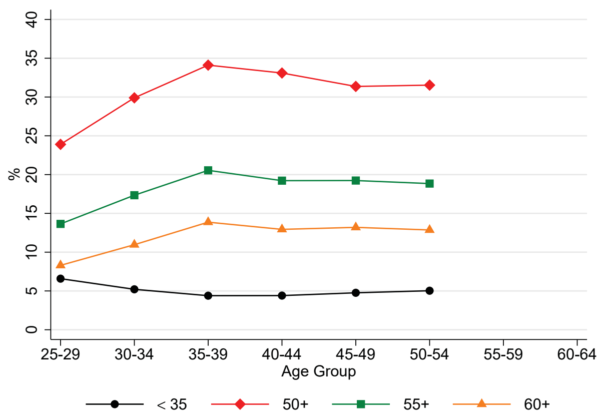


Figure 25: Fraction Working Long Hours by Education

(a) Men (b) Women

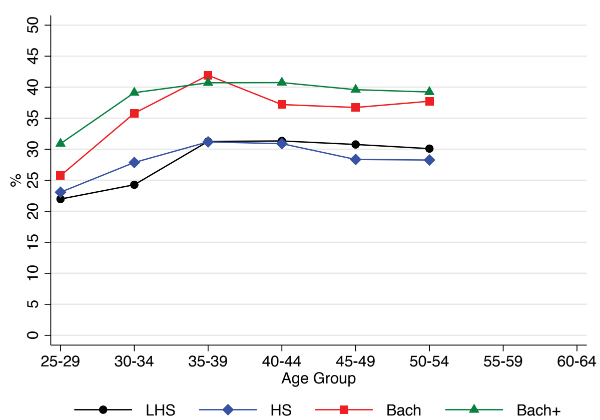


Figure 26: Fraction Working Short Hours by Education

(a) Men (b) Women

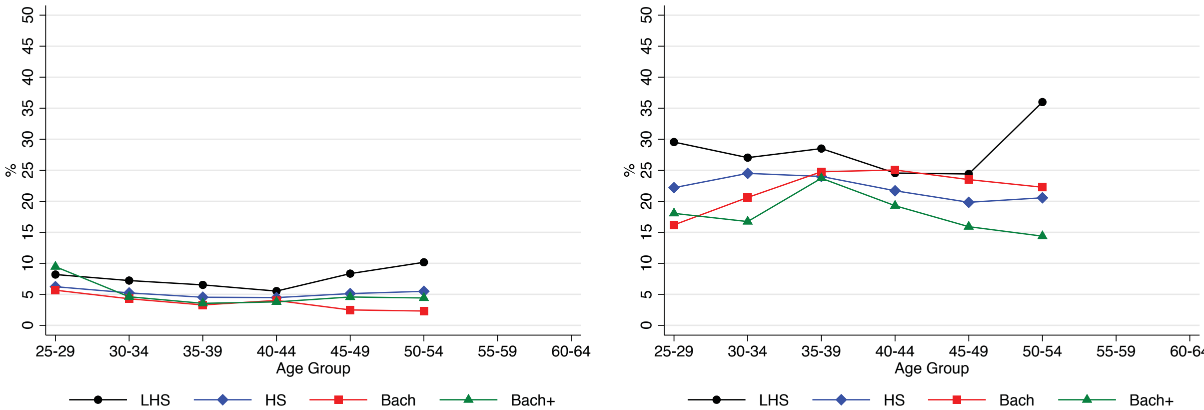


Figure 27: log Hourly Wages, log Annual Earnings and Weekly Hours Worked

(a) log Hourly Wages by Weekly Hours (b) log Annual Earnings by Weekly Hours

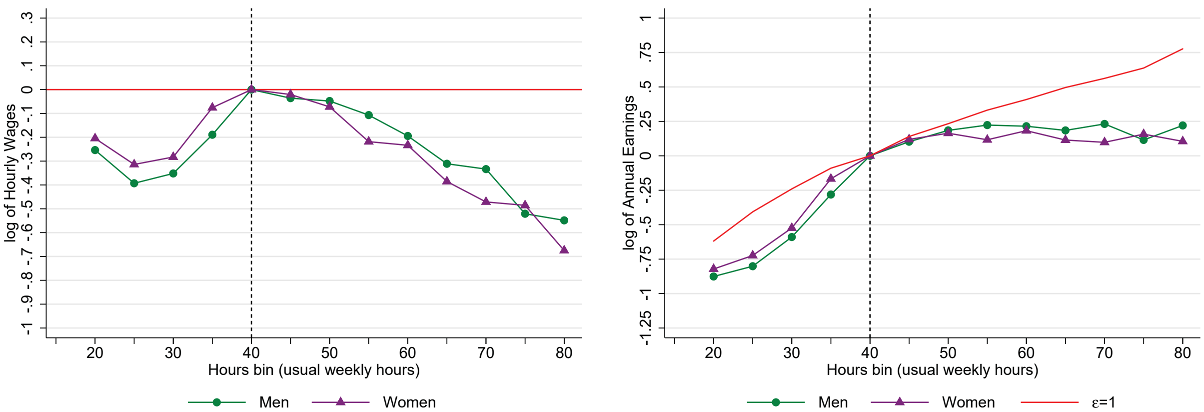


Figure 28: log Hourly Wages and Weekly Hours Worked for Men

(a) By Education (b) By Age