Family Friendly Occupations and the US Gender Wage Gap*

Michael Coelli†

29 October 2012

PRELIMINARY

Abstract

A consistent finding in US labor market research is that the wages of men and women are lower in predominantly female occupations. The roles of a number of specific occupational characteristics that may be of benefit to individuals juggling labor market and child-rearing responsibilities in explaining this relationship are investigated. These occupation level characteristics include the proportion of employees working part-time, the average hours of work among full-time workers and the average commuting time to work. The relationship between average occupation commuting time and wages is examined in detail.

Keywords: gender wage gap, occupational characteristics, commuting time

JEL codes: J16, J24, J31, J71

---

*This research has benefitted from discussions with David Card and Nicole Fortin. Thanks also to seminar participants at the University of Melbourne, UC Berkeley, UC Santa Barbara, University of British Columbia, ESAM Melbourne 2012 and LEW Wollongong 2012. All opinions and any errors are my own.

†Department of Economics, The University of Melbourne. Email: mcoelli@unimelb.edu.au

1 Introduction

Despite some narrowing over recent decades (e.g. Jacobsen (2007), Table 6.4), females and males still predominantly work in different occupations in the US. A Duncan and Duncan (1955) index of occupation dissimilarity using US Census 2000 data over 475 occupations implies that approximately 52% of male (female) employees would have to change occupations to be distributed the same as female (male) workers. It has been well established that the wages of both men and women in the US are lower in occupations where the workforce is predominantly female (see for example Macpherson and Hirsch, 1995). Understanding why this relationship holds is important for our understanding of the labor market, and for evaluating the desirability of comparable worth and other policies aimed at changing the wage structure and the gender wage gap.

A number of potential explanations for differences across gender in the occupations that individuals are employed in have recently been investigated in the economics literature. These include on the job risk of death and injury (DeLeire and Levy, 2004; Grazier and Sloane, 2008), earnings losses after career interruptions (Adda et al, 2010), risk of layoff (Dan, 2010), cross-sectional earnings risk (Bonin et al, 2007), preferences for money (Fortin, 2008), personality and preferences (Borghans et al, 2008; Krueger and Schkade, 2008: Rosenbloom et al, 2008), and non-congitive skills (Fortin, 2008; Antecol and Cobb-Clark, 2010; Cobb-Clark and Tan, 2010).

This paper investigates the role of occupation characteristics that may be grouped under the title of "family friendly" on gender differences in occupations and the gender wage gap. The focus is on: the prevalence of part-time work in the occupation, the average hours of work among full-time workers, and the average commuting time to the job within an occupation. The three attributes of part-time flexibility, lower expectations of long work hours and shorter commutes may be of value to individuals required to balance work and family responsibilities. It is still generally the case that females bear the majority of family and home responsibilities in households (Bureau of Labor Statistics, 2011).

The effect of part-time prevalence in an occupation on earnings and the gender wage gap has been explored previously (Macpherson and Hirsch, 1995). Investigating the role of long hours in occupations on the gender wage gap is a natural extension, and has been discussed as an explanation for occupation differences across genders by Cavallo and O'Neill (2004). The role of average commute times in an occupation has, as far as I am aware, not previously been investigated as a potential contributing factor to the gender wage gap.[1] The main focus of this paper is on this particular occupation attribute. The analysis is conducted primarily with data from the 2000 US Census, a primary source of commuting time information.

To pre-empt the results, females are more likely to have shorter commute times, to work part-time and to work fewer hours even if full-time than males at the individual employee level. In addition, females are more likely to work in occupations that have more family friendly attributes on average: shorter average commutes, higher part-time prevalence and shorter average hours among full-time employees. This holds even when these occupation average characteristics are constructed using averages for male employees only. I find that an individual's own commute time is less positively related to earnings than occupation average commute times.

A positive relationship between occupation average commute times and earnings holds even when allowing for fixed unobserved individual effects in panel estimation.[2] I also employ cross-city variation in commute times to provide support for the contention that it is commute time differences across occupations that is driving earnings differences rather than earnings purely being driving by other occupation characteristics that may be correlated with commute times. The role of additional measures of occupation flexibility (starting time, variation in commute times and variation in hours of work) are also explored, but these measures do not assist in

[1]The factors behind the observed shorter commuting time of females has been a long standing area of research in Urban Studies and Geography. MacDonald and Peters (1996) discusses several contributing factors and surveys the earlier literature.

[2]The panel estimation employs data from the 2004 US Survey of Income and Program Participation (SIPP) panel.

explaining the gender wage gap.

The outline of this paper is as follows. In Section 2, a background discussion of the role of occupation average versus individual commute times in affecting earnings is provided. Estimates relating individual and occupation characteristics with earnings are provided in Section 3. In section 4, details are provided showing that females with children in the home are more likely to work in occupations with "family friendly" characteristics than females without children, but females without children are still much more likely to work in "family friendly" occupations than males. Section 5 concludes.

2 Occupation average versus individual commuting times

The determination of both the living and working arrangements of individuals is a complex process. Individuals choose where to live and work to optimize wellbeing. Individual well-being (or utility) will be a function of earnings from working, non-monetary job attributes, housing costs and house / neighborhood quality. A shorter commute time, other things being equal, would generally be desired.[3] These choices are, however, constrained by opportunities. Housing availability is a function of historical building patterns, zoning regulations, geography and relative demand (reflected in price). Job availability is a function of firm (work) location.

Firms choose where to locate to optimize profits. Some firms may find it more optimal to locate near workers in residential areas (perhaps also to be near customers). Others may optimally choose to locate in city centers (agglomeration spill-overs) or industrial areas, again based on zoning regulations, but also on transportation links. Firm location may also be a function of history, with potentially large re-location costs. As a result, the living and working locations of individuals we observe may or may not be an equilibrium outcome.

[3] Stutzer and Frey (2008) describe the significant negative effect that longer commutes can have on measures of subjective well-being (life satisfaction).

2.1 A simplified model

To illustrate the mechanisms in a world where equilibrium is reached, a simplified model of occupation and location choice in a city is developed in Appendix A. Individuals optimally choose a location specific occupation (city or suburb job) and a suburb to live in (near the city or further out), based on wages paid in different occupations, rental costs in different suburbs, and the individual's heterogeneous distaste for commuting.

Competitive firms can freely locate either in the city or in suburban areas, with potentially different production technologies depending on this location choice i.e. production amenities may differ across locations. Due to free mobility and optimal capital investment decisions, wage differences across city and suburban locations are determined by productivity differences alone. For wages to be higher in city (longer commute) occupations (as we will observe in the empirical analysis to follow), it must be the case that there are productivity benefits for firms locating in city areas, perhaps from spillovers, being closer to suppliers and customers, infrastructure, et cetera. If no such benefits existed, firms would not locate in city centers as they could not compete for workers, given worker distaste for (costs of) commuting.

Rents in suburbs near the city and further away are determined by supply and demand. Supply constraints are determined simply with rents increasing with the number of individuals choosing to live in a particular suburban area. This mechanism proxies the extra costs involved in attempting to add more housing to particular suburbs.

The model in Appendix A allows for multiple equilibria, depending on parameter values. I focus on an equilibrium where individuals live in both inner and outer suburban areas and work in both city and suburban occupations (as we observe). In this equilibrium, wages are higher in city jobs than suburban jobs, while rents are higher in suburbs close to the city relative to suburbs further out, otherwise no individual would choose to live in outer suburbs.[4] Individuals

[4]In this model, there are no amenity differences across suburbs. Some readers may note the low rents in some inner suburb neighborhoods in some cities due to urban decay. Including amenity differences in this model is straightforward, and does not change the basic predictions of the model.

with the lowest distaste for commuting will choose to work in the city and live in outer suburbs, taking advantage of the lower rent in outer suburbs. A second group with distaste values in the middle of the distribution will choose to live in inner suburbs and work in city occupations. The remainder with the highest distaste for commuting will choose to live in outer suburbs and work in suburban occupations.

What do we learn from this model about commuting times, occupations and wages? Earnings differences between long commute occupations (located in city centers) and short commute occupations (located in suburbs close to where people live) are driven by firm technology differences alone. Thus we should not think of there being some "causal effect" of commuting time on earnings. This implies that there is nothing to be gained from searching for "exogenous" variation in commuting times by occupation if the objective is purely to improve empirical estimation of the relationship between occupation commuting times and earnings .[5]

If there was some variation in relative commuting times, i.e. if commuting from outer suburbs to the city became less time-consuming (via an improvement in transport links), then it would change the work location and suburb of residence decisions of some individuals. Some inner suburb individuals will move to the outer suburbs and still commute to city occupations, while a smaller number will move from the outer suburbs where they worked locally to inner suburbs and then will commute to city occupations (in response to inner suburb rent reductions). Overall, there will be more people working in the city but less people living in inner suburbs. This will mean that average commuting times of city workers may not fall much.

The model also illustrates that individuals with the greatest distaste for commuting choose to work in suburban jobs, and in this specific model of households with one worker only, they live in outer suburbs where rents are cheaper. Individual differences in commuting times among workers in the same (city) occupation reflect differences in distaste for commuting i.e. workers

[5]This implication is based, however, on a world with competitive firms with free mobility. If firms are not mobile and competition is not perfect, then "exogenous" variation in commuting times might improve estimation.

with the least distaste for commuting choose to live further from the city and pay lower rents. Wages do not differ by individual commuting distance differences within a (location-specific) occupation, but rents may.[6] Importantly, workers in city occupations are more likely to live in inner suburbs, paying higher rents but earning higher wages. Thus average commuting time differences between city and suburban occupation workers may be small relative to the observed wage premium from working in city occupations, implying large compensating differentials for potentially small average commuting time differences. It must be kept in mind when interpreting the estimates to follow, however, that city workers may be using part of the city occupation wage premium to pay for higher inner suburb rents.

Extensions of the basic model to allow for differences in worker productivity are also considered in Appendix A. For worker productivity differences to alter individual work location and residence decisions, it must be either that productivity and distaste for commuting are correlated, or that productivity differences interact with work location. If productivity and distaste for commuting were negatively correlated, more productive workers would choose to work in city occupations. Also, if there was some complementarity between productivity of workers and city occupations (high productivity workers are relatively more productive in city than suburban occupations), then again, more productive workers would choose to work in city occupations.

If high productivity types are more likely to work in city occupations (due either to complementarity or to a negative correlation between productivity and distaste for commuting), the observed average wage difference between city and suburban occupations will be larger than the wage gain that any specific worker with a particular productivity level would obtain from changing from suburban to city work. In the empirical analysis to follow, the relationship between

[6] Along a similar vein, Timothy and Wheaton (2001) argued that wages in a particular working zone (an area) will be higher if the average commuting time of workers to that zone was higher. Higher average commuting times reflect the need for workers to commute from other zones i.e. the demand for workers in a zone exceeds the supply in the same zone. Individual commuting time differences should not be reflected in earnings differences, but instead should be capitalized into housing price / rent differences, i.e. lower prices for those houses positioned further from employment opportunities.

commuting time and earnings is estimated in a cross-section, and again using panel data with fixed effects. The estimates of the relationship between occupation average commuting time and wages from the fixed effects estimates are somewhat smaller than from the cross-section, suggesting that there may be self-selection of high productivity types into city (long commute) occupations. This is not the only possible explanation of lower estimates when using fixed effects, however, as measurement error may have larger attenuation effects in such estimates.

In the simplest version of the model with frictionless and competitive markets and worker productivity homogeneity, wages differ across occupations but not within them. We will see the significant negative relationship between occupation average commuting times and earnings in the empirical section to follow. There remains, however, a less strong but still negative relationship between individual commuting times and earnings even within occupations. Several explanations for such a relationship at the individual level come to mind, where we relax the assumption of frictionless and competitive markets, and allow for unobserved (to the researcher) differences in worker productivity and firm location / specifics.

- Individual unobserved productivity differences may result in more productive workers who earn more being able to afford longer commutes. Dargay and Ommeren (2005) note the offsetting effects of higher income being able to buy the more expensive properties closer to city centers with the same higher incomes being able to afford the larger homes and sites that are generally available further from city centers. They estimate that higher income causally results in longer commutes.

- Individual firm heterogeneity in location may mean that some firms are located in city centers while others are in the suburbs, and they may employ individuals in the same occupations (e.g. city versus suburban lawyer). Productivity differences in city versus suburban firms may result in wage dispersion within occupations that is related to individual commutes.

- Specific firm / worker interactions (productivity) may mean that some workers are paid more to work in specific firms that are located further from home than another similar firm, and joint household location decisions may restrain the individual from moving closer to that specific firm.

- If there are frictions in labor markets, search models with job posting can create an "implicit" compensating differential (Manning, 2003). In this model, job openings arrive randomly, posted wages are heterogeneous, and individuals with distaste for commutes only accept low wage offers if the commute is shorter.

- Search models with random matching and ex-post bargaining also yield a negative relationship between individual commuting time and wages (Rupert et al, 2009). Individuals have a lower threat point in bargaining if the matched job has a shorter commute. In addition, a potential match is only accepted in the first instance if there is surplus in the match (productivity exceeds the worker's expected alternative).

2.2 Occupation dispersion within cities and commuting time

Average commuting times in an occupation are likely to be lower if jobs in that occupation are more spread among residential areas where people live. Figure 1 presents a scatter plot of occupation average commuting times (the horizontal axis) against a measure of how dispersed the jobs in each occupation are among where people live (vertical axis). This measure of occupation dispersion OD^j in occupation j is based on the Location Quotient (LQ) measure regularly used in geography and urban economics to describe industry concentration in different regions. Here, I define the Location Quotient as follows:

$$LQ_k^j = \frac{\text{\% of workforce in area } k \text{ working in occupation } j}{\text{\% of total national workforce working in occupation } j} \tag{1}$$

An LQ_k^j greater than 1 denotes that occupation j is concentrated in area k. I construct a measure for each occupation combining area-specific LQ_k^j measures using worker residence area weights. The occupation dispersion measure (OD^j) for occupation j can be written as follows, where $k = 1, ..., K$ denotes a specific area and N_k denotes the number of workers who **live** in area k:

$$OD^j = \sum_{k=1}^{K} LQ_k^j \; \frac{N_k}{N} \quad \text{where} \quad N = \sum_{k=1}^{K} N_k \tag{2}$$

High values of OD^j (above 1) will occur in those occupations that are concentrated (high LQ_k^j) in areas where most workers live (are dispersed), as individual area LQ_k^j's are combined using weights based on where workers live.

This occupation dispersion measure was constructed using data from the 2000 US Census 5% micro-data sample, using 1238 separate Census Place of Work Public Use Micro-data Areas (POW PUMAs) and 470 separate Census occupations.[7]

——————— INSERT FIGURE 1 HERE ———————

Note the expected significant negative relationship overall between average commuting times and occupation dispersion in Figure 1. There are, however, a number of occupations lying above and to the right of the general negative relationship. These occupations have relatively high dispersion measures but also longer commuting times on average. These occupations are generally in the construction trades and extraction, such as roofers, pavers, fence erecters, explosives workers, drillers, et cetera. Housing construction in particular will be dispersed among where people live, but due to the contract short-term nature of this kind of work in any one specific location, such workers may choose not to alter their living arrangements for every change in work location.[8]

[7]All military employees and occupations were excluded from the analysis.

[8]Occupation dispersion measures were also constructed using data from the 2000 US Census Transport Planning Package (CTPP). This source provides more detailed geographic breakdowns but only provided occupation details at the two digit level (23 occupation groups). The significant negative relationship between average commuting time and occupation dispersion was even more evident in this data using Census Tracts as areas for California and New York state (the only two states I constructed measures for), with the construction trades and extraction group again lying to the right of the negative relationship. Figures for these relationships are available upon request.

3 Earnings estimates

3.1 Individual characteristics

To investigate the relationships between individual and occupation characteristics and earnings, I employ the two step estimation procedure of Baker and Fortin (1999, 2001). In the procedure's first step, individual log hourly wages are regressed on a set of individual employee characteristics plus a set of indicator variables for each of the 470 individual occupations. In the second step, the estimated coefficients on these occupation indicators are employed as the dependent variable in regressions at the occupation level on a set of occupation level characteristics. The estimates from the first step are presented in this sub-section.

The main argument forwarded by Baker and Fortin (1999, 2001) for employing this two-step procedure rather than the more common one-step method of including the occupation level characteristics directly in the individual log wage regressions (e.g. Macpherson and Hirsch, 1995) is to reduce potential coefficient bias. In the one-step method, if there are missing occupation level characteristics that are correlated with included regressors, the estimates on all coefficients, including the coefficients on the individual level characteristics, may be biased. By including the set of unrestricted occupation indicators in the individual level regressions, the potential for missing occupation level variables is averted, and the estimates on the individual level characteristics should be unbiased. Missing occupation level regressions in the second step may still be a source of bias in the estimates of the coefficients on these occupation level variables.

Summary statistics for the individual level variables included in the first step regressions are provided in Table 1. The data is from the 2000 US Census 5% micro-data sample. Details of sample and variable construction are provided in Appendix B. The gender wage gap is approximately 24% in this data, using constructed hourly wage rates. Females have marginally more years of schooling, and are more likely to be employed by not-for-profit groups and government employers. Regarding individual commuting time, females have on average one way commutes (the Census data collects commute time for the journey to work only) that are shorter by 4 minutes. Females are also more likely to begin their commute to work between the hours of 7 a.m. and 12 noon than males, while males are more likely to start their commutes between midnight and 7 a.m.

——————— INSERT TABLE 1 HERE ———————

The estimates from the first step log hourly wage regressions by gender are provided in Table 2. The estimation results are generally in line with previous research.[9] Education attainment is strongly related to earnings, with education levels below the base category of a high school diploma related to earnings penalties, while post-secondary education is related to significant earnings premia.[10] Immigrants and non-white individuals earn less on average than their US born white counterparts. Employees of not-for-profit groups generally earn less than private sector employees, while federal government employees in particular earn more.

——————— INSERT TABLE 2 HERE ———————

The variables that are generally not included in such models of earnings are the individual commuting time and the time of starting the commute to work. A cubic in commuting time was found to adequately describe the relationships. Figure 2 plots the predicted values from the cubic relationships for females and males. Female earnings rise further with individual commuting time than male earnings, with earnings rising with commute times until commutes reach approximately one hour for both genders. Note that these regressions include unrestricted indicators for 470 Census occupations, for 23 major industry groups and for 2,071 individual Census PUMA geographic locations based on where individuals live. Thus these relationships between commuting times and earnings are over and above the relationships with occupation, industry and living location. Potential explanations for positive relationships between individual commutes

[9]The positive relationship between part-time status (less than 35 hours of work per week) and earnings for males is an exception. This may be due to how hourly earnings was constructed using annual wage and salary income, weeks worked and usual weekly hours, rather than using a direct report of the hourly wage rate.

[10]The relationships are not linear in constructed years of schooling.

and earnings were outlined in the previous section, including unobserved individual and firm effects, and search frictions in labor markets.

——————— INSERT FIGURE 2 HERE ———————

Regarding individual commute starting times, earnings for both genders are higher for those starting their commutes prior to 6 a.m. or from 7 p.m. onwards relative to the base group of starting commutes from 8 to 8:59 a.m. For females, earnings are also higher for those starting between 6 and 7:59 a.m. Earnings for both genders are lower if starting mid-morning (9 to 11:59 a.m.). Thus there may exist compensating differentials for less desirable commute starting times. Earnings are lower for individuals starting in perhaps more family friendly time periods from 8 to 11:59 a.m., which were also the starting times that were more prevalent among female employees.

3.2 Occupation average characteristics

The estimated coefficients on the occupation indicators from the first step are now employed as the dependent variable in the second step estimations at the occupation level. Summary statistics for the occupation level variables included in these second step regressions are provided in Table 3. The first variable is the proportion of females (PF) in the occupation. As expected, the mean of PF for females (0.666) is much higher than it is for males (0.304), given the high level of occupation segregation that remains in the US. The coefficient on PF in wage regressions is a particular focus in the comparable worth literature, including the studies of Macpherson and Hirsch (1995), Baker and Fortin (1999, 2001) and O'Neill (2003).

The next three occupation average variables in Table 3 are the main focus of this research. They were constructed using the Census 2000 micro-data, with the statistics provided based on occupation average characteristics constructed using only male employees. By constructing the averages using male employees only, it avoids inducing a correlation between these three averages and PF. Constructing these averages over both males and females together would by construction result in correlations between these measures and PF, as females are more likely to work part-time, to work less hours even if full-time and to have shorter commutes even within occupations.[11] The means of these occupation average characteristics in Table 3 highlight the higher proportions of females in occupations with characteristics that are more family friendly.

——————— INSERT TABLE 3 HERE ————————

To see the relationships between these three "family friendly" occupation characteristics and PF, scatter plots are provided in Figure 3. The size of the circle for each occupation represents the employment size of that occupation in the US. The lines are non-linear regression lines estimated using those employment weights. In the top panel, several occupations with long commutes can be observed clustered at the lowest end of the PF range. Apart from these occupations, there is a small but steady negative relationship between commutes and PF over the range of PF up to 1. Table A1 in the appendix provides a list of the occupations with the shortest and longest commuting times.

In the second panel of Figure 3, a positive relationship between the proportion of employees working part-time in an occupation and PF can be observed, particularly at higher levels of PF. In the third panel, a negative relationship between average hours of full-time employees and PF is observable, with very few long hours occupations at the highest PF levels.

——————— INSERT FIGURE 3 HERE ————————

The next group of variables in Table 3 - job zone, physical, hazzards, strength and environment - were constructed using information from O*NET, the updated version of the Census Bureau's Dictionary of Occupation Titles (DOT). The construction of these variables is described in Appendix B. Each occupation is allocated to one of five job zones in O*NET, reflecting the amount of preparation (education and training) required for entry into that occupation. Occupa-

[11] Averages over female employees were highly correlated with the averages using male employees. Using averages constructed over male employees only resulted in much more conservative estimates (smaller in absolute value) of the relationships between these occupation average measures and earnings than if averages over both males and females were used, and marginally more conservative estimates than if averages using females only were used.

tions in zone 1 require essentially no or limited preparation, while occupations in zone 5 require extensive preparation.[12] There are only minor differences across genders in the job zones of occupations, with females more likely to be in occupations requiring extensive preparation and males more likely to be in occupations requiring some preparation (zone 2). A higher hazzards measure reflects exposure to particular hazzards on the job. A higher strength measure is allocated to occupations where strength is a required ability. A higher environment measure denotes employment in an occupation that has environmental features that are generally unpleasant. For all three of these occupation based measures, males have higher means.

The next measure - fatal occupation injuries - is based on data reported by the US Bureau of Labor Statistics (BLS) on the numbers of such cases in each occupation.[13] Consistent with previous findings (e.g. DeLeire and Levy, 2004), males are in occupations with considerably higher levels of fatalities.

The last six measures in Table 3 were all taken from the Work Activities component of O*NET, using the Importance (IM) measure. Apart from the computing variable, these measures cover those attributes of occupations that some investigators have suggested females may be more attracted to due to preferences or personality. They include caring and teaching, as well as some related more specifically to interacting directly with people.

Estimates of the relationships between occupation characteristics and log hourly earnings (the second step) are provided in Tables 4 and 5 for females and males respectively. As discussed by Baker and Fortin (2001), there are at least three potential weighting schemes that can be employed when estimating these second step regressions: un-weighted, weighted using the number of observations in each occupation (or more specifically the sum of the person weights in each occupation), or weighted using the estimated variance (inverted) of the occupation indicator coefficients from the first step estimates. The choice between weighted and un-weighted

[12] These job zones are intended to be more broad than the measures included in the original DOT, and essentially replace the Specific Vocational Preparation (SVP) and General Education Development (GED) measures.

[13] Details of variable construction are again in the Appendix.

estimation depends on one's beliefs about the greater potential source of error variance in the second step estimates. Using the estimated occupation coefficients from the first step should result in heteroscedastic errors in the second step estimates (noisier estimates of an occupation's coefficient if it is estimated using a smaller number of observations), thus implying weighted estimation is more appropriate. If, however, the variance of the error in the second step population model is large, un-weighted regression may be preferred. The estimates in Tables 4 and 5 use the estimated variance of the occupation indicator coefficients from the first step as weights.[14]

Four different model estimates at the occupation level are presented in Tables 4 and 5. In model 1, only the proportion female (PF) variable and job zone indicators are included. A negative relationship between hourly earnings and PF is estimated for both genders, with a marginally more negative relationship estimated for females. In model 2, the three "family friendly" characteristics are added to the regressions. These characteristics have the expected signs and are all statistically significant. Earnings are lower in occupations with higher part-time prevalence, lower hours among full-time workers and with shorter average commutes. The coefficients on the average commute variables for both genders are approximately four times the size of the steepest part of the equivalent gender commuting time effects in the individual level regressions of Table 2.[15] Note also that the negative relationship between occupation earnings and PF disappears after adding these three particular variables.

——————— INSERT TABLE 4 HERE ———————

——————— INSERT TABLE 5 HERE ———————

Consistent with findings in the earlier literature (Macpherson and Hirsch, 1995; Baker and Fortin, 2001), the estimated relationship between PF and earnings is sensitive to the choice of

[14]As found by Baker and Fortin (2001), the second step estimates were somewhat sensitive to the weighting scheme employed. Note, however, that estimates of the effect of occupation average commute times on earnings - the main focus of this analysis - were found to be very similar across weighting schemes. Estimates using the alternative weighting schemes are available upon request.

[15]The slopes of the cubic commute time functions are steepest at a zero commuting time, and at this point the slope equals the coefficient on commute time in levels.

other variables included in the estimated models. Model 3 includes a set of occupation level variables (described above) that have either commonly been included in such models in the previous literature, or have been discussed as potentially affecting earnings differences across occupations. Note that the estimated effect of PF on earnings is much more negative than in model 1. These additional variables do not all have the expected relationship with earnings. While occupations in higher job zones (more required preparation) have higher earnings, occupations with higher strength requirements and poorer working environments appear to earn less. Occupations with higher exposure to hazzards, however, have higher earnings, as expected. The relationship of earnings with fatal injuries has an unexpected negative sign, but is imprecisely estimated. Note that the hazzards, strength and environment variables in particular are highly correlated with each other. The correlation coefficients among all the occupation level variables are presented in Table 6.

——————— INSERT TABLE 6 HERE ———————

Model 4 includes both the three "family friendly" occupation characteristics and the set of additional occupation level variables of model 3. Inclusion of the three "family friendly" variables lowers the size of the negative relationship between PF and earnings considerably, thus these three variables appear to account for a significant portion of the lower earnings in female-dominated occupations. Inclusion of the additional set of occupation level variables does generally result in smaller estimated effects of the three "family friendly" characteristics on earnings (model 4 versus model 2), but the effects retain their statistical significance, and the commuting time effect falls only slightly for females.

To gain some understanding of how much of the gender wage gap can be attributed to differences across genders in these occupation level characteristics, some simple Oaxaca (1973) and Blinder (1973) decompositions were constructed using the estimates of Tables 2, 4 and 5. In Table 7, a standard decomposition using the estimates from the first step regressions of Table 2 is presented. Individual level characteristics only are included here. The overall log hourly wage gap is 0.242. None of the individual level characteristics contribute much to the gender wage gap. Education differences contribute negatively to the gap, as females have on average higher levels of education. The Census does not have any measures of actual work experience of individuals, and the potential experience measure employed here differs very little across genders.[16] Differences in employer only contribute a small amount to the gap. Females are more likely to work for low paying not-for-profit employers than males, but more likely to work for somewhat higher paying state and local government employers. Individual commuting time differences contribute a small amount to the gap (2%), while starting time differences contribute 2.5%. There is a large estimated contribution of industry differences (12.4%) to the gender wage gap, with males more likely to work in industries that pay higher wages (construction, utilities, manufacturing).[17]

——————— INSERT TABLE 7 HERE ———————

Table 8 presents Oaxaca-Blinder decomposition estimates based on the occupation level estimates of Tables 4 and 5. The percentage figures are calculated as a percentage of the overall gender log hourly wage gap of 0.242. Focusing on model 4, differences in the "family friendly" characteristics all contribute significant amounts to the gender wage gap. The contribution of occupation average commute differences, at 6.2%, is over three times the size of the contribution of individual commute differences within occupation reported in Table 7. Job zone differences subtract from the gap, as females are marginally more likely to work in occupations requiring more preparation. Differences in hazzards, strength requirements and poor work environments also subtract from the gap, due to the negative coefficients on these variables for all but hazzards. While fatal injury differences across genders also appear to subtract from the gap, the estimate is based on coefficients that are imprecisely estimated.

——————— INSERT TABLE 8 HERE ———————

[16]O'Neill (2003) finds that actual work experience differences can account for a significant portion of the gender wage gap.

[17]All estimation results for models that do not control for major industry are available upon request.

3.3 Cross-city variation

The above estimates of the positive relationship between occupation average commute times and earnings may be biased due to missing characteristics of occupations that are correlated with average commuting times (firm location in a city). In this sub-section, I employ cross-city variation in commuting times by occupation to provide further evidence that commuting time differences in particular are related to occupation earnings differences.

Average commuting times vary considerably across cities in the US. The top left hand panel of Figure 4 plots average commuting times in a city against city size (in thousands of employees). The 13 diamond markers in the plot represent the 13 largest cities (Consolidated Metropolitan Statistical Areas - CMSAs) in the US. In decreasing size these cities are: New York, Los Angeles, Chicago, Washington DC, San Francisco (Bay Area), Philadelphia, Dallas, Boston, Detroit, Houston, Atlanta, Seattle and Minneapolis. The square marker labeled "medium" is the average commute across 38 medium sized US cities (MSAs), with the number of employees in these cities ranging from approximately 200 to 800 thousand. The square marker labeled "small" refers to the remaining 192 cities (MSAs) identified in the Census micro-data, with the number of employees below approximately 200 thousand.[18]

———————— INSERT FIGURE 4 HERE ————————

The top right panel of Figure 4 plots the standard deviation of average occupation commute times within each city against city size. For example, the point for New York at the right of the plot denotes the standard deviation of occupation average commute times in New York is nearly 5.5 minutes. Occupation average commute times were constructed separately for each city or city group (small and medium). To avoid small cell sizes, these occupation average commutes by city were constructed for 93 three digit occupation groups, rather than separately by the

[18]Note that the calculation of number of employees by city was based on the person weights provided in the US Census micro data used in the main analysis. Thus these numbers of employees will understate the true number as sample observations were excluded for a number of reasons. See Appendix B for details.

470 occupations identified in the Census micro-data. Note the positive relationship between these standard deviations and city size. This tells us that the dispersion in travel times across occupations is larger in large cities. Occupations that are dispersed among the suburbs where people reside have low commuting times that do not vary too much with city size. On the other hand, occupations that are more centralized in financial district or industrial areas have longer commutes in larger cities relative to smaller cities.

The bottom left panel of Figure 4 also illustrates the "spreading out" effect of city size on occupation average commutes. The slope coefficients from simple regressions of each city's occupation average commutes versus occupation average commutes in medium cities are plotted here. Occupation average commutes in medium cities are used as the regressor in each case. The slope coefficient using medium cities themselves as the dependent variable is of course equal to one. These coefficients generally rise above one as city size increases. To see what is going on in more detail, occupation average commutes for Los Angeles (LA) are plotted against occupation average commutes in medium cities in the bottom right panel of Figure 4. Commuting times by occupation generally lie in a line above the 45 degree line, with points very close to the line for low commute occupations. Thus local occupations have equally low commutes in large LA and medium sized cities, while jobs that are on average more distanced from residential areas have relatively longer commutes in larger LA. The slope of the regression line here of 1.36 is what is plotted for LA in the bottom left panel of Figure 4.

I now employ this cross-city variation in commuting times by occupation to observe whether occupation average commuting time differences are related to earnings differences, holding constant other occupation characteristics that may be correlated with commuting times. To do this, I estimate individual level log hourly wage regressions essentially identical to those reported in Table 2, but now also adding a variable of commuting time by occupation within each city. Employees residing in non-metropolitan or mixed metropolitan / non-metropolitan areas were excluded from these regressions. Note that these regressions include the full set of unrestricted occupation indicators and individual city (CMSA or MSA) indicators.

The coefficients of interest from these regressions are those on the added commuting time by occupation and city (or city group) variable. For females, the coefficient equalled 0.0057 (t-statistic of 5.77), while for males it equalled 0.0045 (t-statistic of 5.20). Note that these t-statistics were constructed using standard errors that allowed for clustering at the city by occupation level. These estimated relationships are significant and relatively large, although they are only half the size of the estimates from the occupation level regressions of Tables 4 and 5. This may be due to measurement error in the city by occupation commute time variables, attenuating coefficient estimates towards zero. Constructing commuting times separately by occupation and city resulted in averages being constructed with only a small number of observations in some cases.

Here we have observed that differences in earnings between suburban and city occupations (short and long commutes) are larger in larger cities, where the differences in commuting times between city and suburban occupations are also larger. We must take care, however, not to interpret this as reflecting a causal effect of occupation commuting times on earnings. Recalling the theoretical model described above, earnings differences by occupation are driven by firm location productivity differences alone, given mobility (free entry and exit) of competitive firms. This finding thus suggests that firms located in city centres are more productive in larger cities, which in turn suggests that agglomeration externalities may be larger in larger cities.

3.4 Self-selection - fixed effects estimates

A common concern in studies of occupation and earnings is that the lower average earnings in certain occupations may be due to selection effects based on an unobservable component of an individual's productivity. For example, earnings may be lower in food preparation occupations that are located in residential areas than in legal occupations that are predominantly located in city centers. The earnings difference may reflect the lower productivity of those individuals in food preparation relative to law. The discussion of the extension to the theoretical model including heterogeneity in worker productivity also illustrated that positive self-selection may occur if worker ability was relatively more productive in city (long commute) occupations, or if ability and distaste for commuting were negatively correlated.

The standard procedure for dealing with such selection concerns if some aspects of ability are not measured is to estimate earnings models using longitudinal data and controlling for a fixed unobserved individual effect. This procedure will control for any time-invariant productivity differences across individuals. The estimated effects of occupation characteristics on earnings are then identified via individuals that change occupations over time.

I employed data from the 2004 Survey of Income and Program Participation (SIPP) to estimate panel data models with individual fixed effects. In the SIPP, survey respondents are interviewed 3 times a year for up to four years. Information on employment and earnings is collected each survey wave. I attempted to estimate earnings models that were as close as possible to those estimated using Census data, with a few differences. One notable difference is the use of direct reports of the hourly wage rate, when it was provided, rather than relying on the constructed measure from the Census. There is also no individual commuting time or starting time information in the SIPP, so these variables were unable to be included. Details of SIPP sample and variable construction are provided in Appendix B.

Estimates from both pooled and fixed effect log hourly wage regressions using the SIPP data are presented for females and males in Tables 9 and 10 respectively. The pooled estimates are provided for comparison purposes, both to the fixed effects estimates and to the estimates using Census data. Note that a one step estimation strategy was employed here, rather than the two step strategy employed using the Census data. Occupation level characteristics were entered directly into the individual log hourly wage regressions. The one step estimation strategy is more amenable to fixed effects estimation.

——————— INSERT TABLE 9 HERE ———————

——————— INSERT TABLE 10 HERE ———————

The estimated coefficients on the individual characteristics from the pooled regressions are quite similar to those estimated using Census data. One exception is the negative coefficients on the part-time indicator estimated using the SIPP. This is the more standard finding in the literature i.e. lower wage rates for part-time workers. The use of a constructed hourly wage in the Census data may be the source of the zero (females) and positive (males) part-time effects estimated. The estimated coefficients on the occupation characteristics are also generally in line with those from model 4 using Census data. Note that the occupation level measures for the three "family friendly" characteristics are the ones constructed using the Census data, and linked to individuals in the SIPP data via the occupation individuals reported working in.

Turning now to the more important fixed effects estimates, the coefficients are generally attenuated towards zero relative to the pooled estimates. Again a notable exception is the coefficient on the individual level part-time indicator, which is positive rather than negative for both genders. Regarding the "family friendly" occupation level characteristics, the estimates for part-time prevalence generally maintain their size and statistical significance. The coefficients on average hours if full-time are no longer statistically significant, while for average commutes, the coefficients are smaller but are still statistically significant. Thus the estimated commuting time effects do not purely reflect selection on ability. The attenuated size of the estimates may be due in part to compounding measurement error in occupation reports that arise in fixed effects estimates.

These fixed effects estimates of the commuting time relationship with earnings are essentially the same for females and males (0.0041 and 0.0034 to the fourth decimal point respectively). In the cross-section, the relationship for females was larger. The theoretical model extended to the case of heterogeneous ability implied that if females had the highest distaste for commuting, then any self-selection bias in the cross-section should be larger for females. These estimates are thus consistent with this model implication.

If self-selection bias in the estimates of the relationship between occupation commutes and earnings is present in the cross-sectional estimates, as the fixed effects estimates suggest, it implies that the earnings gain that any individual may obtain from moving from a suburban to city occupation (short to long commute) will be less than the cross-section estimates imply. It may be the case, however, that the fixed effects estimates may actually understate the average earnings gain across the ability distribution, if we only observe low ability types switching in the data. We may observe more low ability types switching if they are more likely to be at the margin of the city / suburb occupation choice. High ability types may always be observed in city occupations if ability is more productive in such occupations.

3.5 Additional measures of occupation flexibility

The focus above has been on three particular "family friendly" characteristics of occupations that may be related to lower equilibrium earnings in those occupations. There may be additional characteristics of occupations related to flexibility of work arrangements that may also be attractive to individuals who need to balance work and family responsibilities. Two additional measures investigated here are based on variation in commuting times and variation in hours of work if full-time i.e. the second moments of these two variables. For example, individuals may still choose to work in occupations with long commutes or long hours on average if there are some jobs within such occupations that have short commutes and short hours. Specifically, measures of the coefficient of variation (standard deviation divided by the mean - again using male employees only) within occupations in commuting time and hours if full-time are investigated here. Summary statistics for these two occupation level coefficients of variation are provided in Table 11. The means across individuals are the same for females and males for variation in hours if full-time (0.170), and are similar for variation in commutes (0.887 and 0.889).

——————— INSERT TABLE 11 HERE ———————

Individuals may also value occupations with variation in home leaving times, as it may provide the flexibility to juggle multiple responsibilities. Additional occupation level variables were constructed denoting the proportion of employees in an occupation (again using male employees only) with leaving times within certain ranges of starting times. Summary statistics for these occupation level variables are also provided in Table 11. These measures do differ across genders. For example, males are more likely to work in occupations where a higher proportion of employees leave for work in the two early morning starting time ranges: between midnight and 5:59 a.m. and between 6 and 6:59 a.m. Females are more likely to be in occupations where higher proportions of (male) employees leave for work in the middle four time categories: from 7 a.m. until 6:59 p.m.

Finally, a measure reflecting the level of variation in home leaving times within each occupation was constructed as the sum of the squares of the proportions within each starting time range. Summary statistics for this variable are provided in the last row of Table 11. Higher values for this variable denote occupations where starting times are inflexible i.e. most employees start in one or a small number of time ranges. Note that the means for this measure are very similar for males and females.[19] It appears that males are in jobs with starting times concentrated earlier in the day, while females are in occupations with starting times concentrated later in the day.

Occupation level regression estimates including these additional measures of occupation flexibility are presented in Table 12. Higher commute and hours if full-time coefficients of variation are related to lower earnings, although the hours variation is not statistically significant for males. Thus it appears that earnings are lower in occupations offering greater flexibility along these two specific dimensions. Regarding occupation leaving times, only very early or very late leaving times are statistically related to earnings. What was not expected is that occupations with a greater proportion of male employees leaving home very early in the morning pay less

[19] Golden (2001) and McCrate (2005) investigated differences across genders in self-reports of start and end time flexibility in the job i.e. whether individuals had the opportunity to change their work schedules. They found that males were slightly more likely to have such flexibility rather than females, and McCrate (2005) found that earnings were not necessarily lower in jobs with such flexibility.

than other occupations rather than more. Finally, occupation concentration in leaving times is positively related to earnings. Earnings thus seem to be higher in occupations with home leaving times that are more concentrated within specific parts of the day (less flexible). Higher earnings are required to induce individuals to work in such occupations.

——————— INSERT TABLE 12 HERE ————————

Even though these additional measures of occupation flexibility were found to be related to earnings, these measures cannot assist us in understanding the gender wage gap. As observed in Table 11, these three measures had essentially the same means across males and females. Thus it does not appear as if females are more responsive to occupation flexibility than males, at least along the three dimensions investigated here.

4 Gender, children and occupation attributes

The summary statistics in Table 3 illustrated that females are more likely than males to be working in occupations with certain "family friendly" attributes: shorter average commutes, more prevalent part-time work and shorter average hours if full-time. The regression results above revealed that such occupations also have lower average earnings. If some females are trading off such "family friendly" characteristics for lower earnings, it should be those with more family responsibilities (children in the home) that are doing the trading. To investigate this issue, employees were separated into groups based on age, gender, marital status and presence of children in the home by child age group. For each of these groups, averages were constructed for each of the three main "family friendly" occupation characteristics based on the occupation the individual was employed in. These averages are presented in Figures 5a to 5c.

In Figure 5a, means for each group of the part-time prevalence of the occupation are presented. Note the strong U-shaped pattern with age for both genders and irrespective of the presence of children. Both young employees (aged 16 to 24) and old employees (aged 55 to 64) are more likely to be employed in occupations with a high part-time prevalence. For females, the presence of children of school age in the home (aged 6 to 17) is also related to being employed in occupations with higher part-time prevalence. Older females with only young children in the home actually are less likely to be in such occupations than those with no children. There is little difference among females by marital status. For males, on the other hand, marital status is related to the part-time prevalence of the occupation, with married males less likely to be in occupations with high part-time prevalence. The presence of children in the home is generally unrelated to part-time prevalence for males. Finally, regardless of the presence of children or marital status, females are much more likely to be in occupations with high part-time prevalence than males.

——————— INSERT FIGURE 5 HERE ———————

In Figure 5b, means for each group of the hours if full-time in the occupation are presented. Here there is generally an inverted U-shape with age for females, while for males this measure first rises with age then flattens out at older ages. For younger females, the presence of children is related to lower average occupation hours, but this is reversed for older females. Again, the presence of school-aged children is related to lower full-time hour occupations. For males, the presence of children is again less important than marital status. Married males are more likely to be in occupations with long hours if full-time. Once more, irrespective of the presence of children or marital status, females are much more likely to be in occupations with fewer hours if full-time than males.

In Figure 5c, means for each group of occupation average commuting times are presented. Inverted U-shaped patterns are again present, with both the young and old more likely to be in short commute occupations. Similar patterns with presence of children and marital status as those described above appear here. Presence of school-aged children matters more for females, while marital status matters more for males. Yet again, irrespective of the presence of children or marital status, females are much more likely to be in occupations with shorter commutes than males.

For females, it appears that the presence of school-age children in the home is a strong driving force behind working in a "family friendly" occupation. The need to get children off to school and to pick them up afterwards is likely to be the source of this finding. For males, the presence of a spouse may allow them to work in occupations that are less "family friendly" but pay higher wages on average. There is also the possibility of a reverse causality here, i.e. males in higher paying occupations are more likely to be married.

5 Conclusions

The analysis above illustrated the negative relationship between three "family friendly" occupation characteristics and earnings. The finding that females were much more likely to be working in occupations that are "family friendly" can assist us in understanding the observed relationship between occupation differences across genders (the proportion female effect) and the gender wage gap.

There are, however a number of questions that remain to be answered.

(1) Why are females more likely to be working in "family friendly" occupations even if no children are present in the home? Are females choosing "family friendly" occupations for other reasons (caring jobs, interactions with people)? Were females "crowded" into these occupations due to segregation? Did they enter such occupations in expectation of potential family responsibilities, either to care for their own children or for other family members?

(2) Do females change occupations towards "female-friendly" occupations once family responsibilities arrive, or are they working in such occupations prior to this? Research by Connolly and Gregory (2008, 2009, 2010) finds that UK women do seem to shift to occupations with more part-time opportunities after having children. Do US employees do the same regarding part-time prevalence, hours if full-time and commuting time?

(3) Have those occupations that females tend to work in for other reasons ("caring" roles, interactive jobs) developed "family-friendly" attributes in response to the desire of the majority female workers to have such job attributes?

There is thus still work to do to fully understand the relationships identified here.

A Theoretical Model

Consider the following simplified equilibrium model of occupation and location choice in a city. Individuals optimally choose a location specific occupation and a suburb to live in according to the following linearized indirect utility function, where w_j is the market equilibrium wage in occupation j, r_k is equilibrium rent in suburb k, t_i is a parameter reflecting individual i's distaste for commuting, while d_{jk} is the commuting time from suburb k to the location of occupation j.

$$U_{ijk} = w_j - r_k - t_i \cdot d_{jk} \tag{3}$$

Individual distaste for commuting t_i is heterogeneous across the population, and is assumed to be distributed $U[T_l, T_u]$, where $T = T_u - T_l$ and $T_u > T_l > 0$. We could think of T_l as the monetary cost of commuting per unit of time, while individuals have some non-zero individual distaste for commuting over and above the monetary cost. The total number of individuals in the city is normalized to equal 1, and they supply labor inelastically. Individuals choose between two occupations: a city occupation paying w_c (e.g. lawyer) and a suburban occupation paying w_s (e.g. supermarket checkout operator). Individuals also choose between two suburb types: suburbs near the city with rent r_n (inner or near) and suburbs further away with rent r_o (outer). If an individual chooses a suburban occupation, the commuting distance d_{jk} is assumed to be zero irrespective of the suburb lived in ($d_{sn} = d_{so} = 0$). If a city occupation is chosen, $d_{cn} = 1$ while $d_{co} = 1 + \alpha$ where $\alpha > 0$, i.e. outer suburbs are more distant from the city that inner (near) suburbs by the factor α.

Firms in this city are competitive and produce a composite consumption good Y according to the following standard constant returns to scale Cobb-Douglas production function.[20]

$$Y_j = X_j \cdot N_j^{\beta} \cdot K_j^{1-\beta} \tag{4}$$

Competitive firms can freely locate either in the city or in suburban areas, with the technology parameter X_j governing production potentially differing depending on this location choice i.e. production amenities may differ across locations. Output is assumed tradeable across cities and even between the city and suburbs,[21] with the price determined externally to any city (price is normalized to equal 1). Firms can rent capital K on competitive markets at p_k per unit.

Firm profit maximization results in an equilibrium wage in each occupation as follows.

$$w_j = \frac{\partial Y_j}{\partial N_j} = \beta(1-\beta)^{(1/\beta-1)} p_k^{(1-1/\beta)} \cdot X_j^{1/\beta} = \Omega(\beta, p_k) \cdot X_j^{1/\beta} \tag{5}$$

Due to free mobility and optimal setting of capital, wages are pinned down by production function parameters and the cost of capital. For wages to be higher in city (longer commute) occupations (as we will observe in the empirical analysis to follow), it must be the case that there are productivity benefits ($X_c > X_s$) for firms locating in city areas, perhaps from spillovers, being closer to suppliers and customers, infrastructure, et cetera. If no such benefits existed, firms would not locate in the center of cities as they could not compete for workers, given worker distaste for (costs of) commuting.

To keep the model as simple as possible, I follow Moretti (2011) by assuming the following equation for rent determination in the two suburbs.

$$r_k = z + \rho_k \cdot N_k \qquad \text{where } \ \rho_k \geq 0 \quad k = n, o \tag{6}$$

[20]The Cobb-Douglas production function is employed for illustrative purposes only. The main tenor of the results will follow in more general constant returns to scale production functions.

[21]Output of city and suburban firms being different to each other but still tradeable across cities would not change the model predictions to any important extent. Price differences across the goods would change equilibrium wages in each occupation, akin to a change in X_j.

This rental determination equation has rents increasing with the number of individuals choosing to live in a particular suburban area. This mechanism proxies the extra costs involved in attempting to add more housing to particular suburbs. It would seem appropriate for $\rho_n > \rho_o$, i.e. that it is more costly to squeeze more housing into inner suburbs that are constrained from expanding relative to outer suburbs, where it may be possible to expand outside current residential boundaries. If ρ_n were infinite, for example, the number of houses in inner suburbs would essentially be fixed at some level N_n. Lower values of ρ_k reflect more elastic supply of housing.

In this model, different model parameters will yield different equilibria. If parameters were such that $w_c < w_s + T_l$, no individual would choose to work in the city, as wages would not be high enough to overcome positive commuting costs. If $w_c > w_s + (1+\alpha) \cdot T_u$, then all individuals would choose to work in the city. To generate equilibria where individuals work in both the city and suburban areas, we will focus on parameter values (values of $X_c, X_s, \alpha, \beta, p_k, T_l$ and T_u) such that $T_l < w_c - w_s < (1 + \alpha) \cdot T_u$.

Consider an equilibrium where individuals live in both inner and outer suburban areas and work in both city and suburban occupations. The parameter values that lead to such an equilibrium will be discussed below. In such an equilibrium, where necessarily $r_n > r_o$ (otherwise no individual would choose to live in outer suburbs), individuals will choose their optimal living and working arrangements depending on their individual commuting distaste parameter t_i.[22] Those with commuting distaste values in a range $T_l \leq t_i < T_1$ will choose to work in the city and live in outer suburbs, as they have the least distaste for commuting, and will take advantage of the lower rent in outer suburbs.[23] A second group with distaste values in a range $T_1 \leq t_i < T_2$ will choose

[22]In this model, there are no amenity differences A_k across suburbs. Some readers may note the low rents in some inner suburb neighborhoods in some cities due to urban decay. Including amenity differences in this model is straightforward, and if preferences for such amenities were homogeneous across the population, individuals would optimize over $r_k - A_k$ rather than just r_k. In this case, r_o may be less than r_n in equilibrium if $A_n < A_o$, but $r_n - A_n$ will be greater than $r_o - A_o$.

[23]There will be a positive number of individuals in this category as long as $r_n - r_o > \alpha T_l$, i.e. as long as there is a group with low distastes for commuting, or housing supply is such that rents are a certain amount higher in inner versus outer suburbs.

to live in inner suburbs and work in city occupations. The remainder with the highest distaste values in the range $T_2 \le t_i \le T_u$ will choose to live in outer suburbs and work in suburban occupations. No individuals choose to live in inner suburbs and work in suburban occupations in this simplified model, as wages are the same as in outer suburb jobs while rent is higher.[24]

To describe the equilibrium here, we solve for T_1 and T_2, the values of t_i where individuals are indifferent between adjacent choices. At T_1, individuals working in the city are indifferent between living in inner versus outer suburbs, thus equation (7) must hold.

$$w_c - r_n - T_1 = w_c - r_o - (1+\alpha)\cdot T_1 \qquad \Rightarrow \qquad r_n - r_o = \alpha \cdot T_1 \tag{7}$$

Note also for this low t_i group, individuals must prefer working in a city versus suburban occupation while living in an outer suburb. For this condition to hold, it requires:

$$w_c - r_o - (1+\alpha)\cdot t_i > w_s - r_o \qquad \Rightarrow \qquad t_i < (w_c - w_s)/(1+\alpha) \tag{8}$$

The value for T_1 constructed using condition (7) ensures that condition (8) holds i.e. (7) is the binding constraint here (see equation (12) below).

At T_2, individuals are indifferent between working in the city and living in inner suburbs versus living and working in outer suburbs. Thus equation (9) must hold.

$$w_c - r_n - T_2 = w_s - r_o \qquad \Rightarrow \qquad r_n - r_o = (w_c - w_s) - T_2 \tag{9}$$

Using the assumed uniform distribution for t_i, we can write the following for N_{cn}, the number of individuals choosing to live in inner suburbs and work in the city. Recall that the total number of individuals in the city has been normalized to 1.

$$N_{cn} = \int_{T_1}^{T_2} f(t_i)\, dt_i = \frac{T_2 - T_1}{T} \qquad \Rightarrow \qquad T_2 - T_1 = N_{cn} \cdot T \tag{10}$$

[24]Certain possible extensions to the model that may yield equilibria where individuals choose to live in inner suburbs but work in suburban jobs include: (a) heterogeneous preferences for inner suburb living, (b) joint household location decisions where one household member may choose to work in the city and another member in a suburban occupation, and (c) certain output produced in the suburbs is non-tradeable or costly to trade.

We can combine equations (7), (9), (10) with equations determining rent determination (6) to solve for the following:

$$N_{cn} = \frac{(1+\alpha)\rho_o \ + \ \alpha\,(w_c - w_s)}{(1+\alpha)(\rho_n + \rho_o) \ + \ \alpha T} \tag{11}$$

$$T_1 = \frac{1}{1+\alpha}\,[(w_c - w_s) \ - \ N_{cn} \cdot T] \tag{12}$$

$$T_2 = \frac{1}{1+\alpha}\,[(w_c - w_s) \ + \ \alpha \cdot N_{cn} \cdot T] \tag{13}$$

The following comparative statics for N_{cn} - the number of individuals living in inner suburbs (and working in the city) - are in line with expectations:

1. N_{cn} falls with ρ_n i.e. as the supply of inner suburb housing becomes more inelastic.

2. N_{cn} rises with ρ_o i.e. as the supply of outer suburban housing becomes more inelastic.

3. N_{cn} rises with $(w_c - w_s)$, the wage gap, as city occupations pay a larger premium.

4. N_{cn} falls with T, the variation in distaste for commuting.

5. N_{cn} rises with α, the extra amount outer suburb dwellers must commute to work in the city.

The comparative statics for $N_o = N_{co} + N_{so}$, the total number of people living in outer suburbs irrespective of where they work, are the opposite of the above. The solutions for N_{co} and N_{so} - the two components of N_o - are as follows.

$$N_{co} = \left[\frac{(w_c - w_s)(\rho_n + \rho_o)/T - \rho_o}{(1+\alpha)(\rho_n + \rho_o) \ + \ \alpha\, T}\right] \ - \ \frac{T_l}{T} \tag{14}$$

$$N_{so} = \frac{T_u}{T} \ - \ \left[\frac{(w_c - w_s)[(\rho_n + \rho_o)/T \ + \ \alpha] \ + \ \alpha\rho_o]}{(1+\alpha)\,(\rho_n + \rho_o) \ + \ \alpha\, T}\right] \tag{15}$$

While N_{co} rises with the wage gap $(w_c - w_s)$, N_{so} falls. It can also be shown that N_{co} falls as the extra commuting time from outer suburbs α increases, while N_{so} rises. Intuitively, N_{so}

and N_{co} rise with ρ_n and fall with ρ_o.[25] The total number of individuals working in the city is $N_c = N_{cn} + N_{co} = 1 - N_{so}$. Thus N_c will have the opposite comparative statics to N_{so}.

We can also solve for the parameter values that ensure $r_n > r_o$, the requirement for this particular equilibrium to hold:

$$\frac{\rho_n}{\rho_o} > \frac{T}{(w_c - w_s)} - 1 \tag{16}$$

The equilibrium thus requires the elasticity of inner suburb housing supply to be relatively low, and it must decrease as the wage gap increases. In the case where ρ_o equals zero i.e. the supply of housing in outer suburbs is perfectly elastic, the condition collapses to $w_c - w_s > 0$. In the case where ρ_n is infinite, and thus the supply of inner suburb housing is fixed at some level N_n^*, the required condition is $N_n^* < (w_c - w_s)/T$.

A.1 Model extension - worker productivity differences

Let us now assume that there are two types of workers in the economy, with high and low ability (productivity). To begin, we will assume that both types have exactly the same distribution of the distaste for commuting parameter t_i. If we assume that the high ability types are equally more productive than low ability types in city and suburban jobs, we have $w_c^h - w_s^h = w_c^l - w_s^l > 0$ where $w_c^h > w_c^l$ and $w_s^h > w_s^l$. In this specific case, the points of indifference between adjacent decisions in equations (7) and (9) are the same for high and low ability types ($T_1^h = T_1^l$ and $T_2^h = T_2^l$). Thus adding worker productivity heterogeneity to the model in this specific manner will not change the model predictions in any interesting way.

A more interesting case is where there is complementarity between ability and working in city occupations i.e. high ability types are relatively more productive in city versus suburban jobs. The city versus suburban occupation wage premium may be larger for high ability workers relative to low ability workers ($w_c^h - w_s^h > w_c^l - w_s^l$ and again $w_c^h > w_c^l$ and $w_s^h \geq w_s^l$). In

[25]These populations will fall with ρ_o as long as $(w_c - w_s) < T + (1 + \alpha)\rho_n$, i.e. as long as the wage gap is not too large. This same requirement must also hold for there to be a positive number of people living in outer suburbs.

this case, $T_1^h = T_1^l$ but $T_2^h > T_2^l$ (rents are the same for low and high ability workers living in the same suburb). High ability workers are more likely to choose to work in city occupations and live in inner suburbs (can afford the higher rents) than low ability workers. However, high ability workers are no more likely to work in city occupations and live in outer suburbs than low ability workers, as the relevant tradeoff (binding equation (7)) is based on rent differences only and not on city-suburb wage premia. Overall, high ability workers are more likely to work in city occupations, and low ability workers are more likely to live in outer suburbs, irrespective of where they work.

If there is complementarity between ability and working in city occupations as defined above, the higher average earnings of city workers may in part reflect higher ability, as these types are more likely to choose to work in city occupations. Thus the raw average wage difference between city and suburb occupations may reflect in part the higher ability of city workers. If females have the highest distaste for commuting, only high ability females will choose to work in city occupations. A mixture of high and low ability males may choose to work in city occupations. The observed mean wage difference between city and suburban occupations among females is thus likely to be larger than that for males. This is precisely what we observe in the empirical analysis, with a larger estimated effect of average occupation commuting time on female wages than male wages.

A further potential extension to the model would be to allow for a correlation between individual ability and distaste for commuting. There is no inherent reason to suspect that such a correlation exists, but some may think that females (with potentially higher distaste for commuting) may be less productive than males in potentially competitive city occupations. Let us assume that high ability types do have lower distaste for commuting. Even without complementarity between ability and city occupations, high ability types would then be more likely to work in city occupations.

B Data construction

The US Census 2000 sample employed in estimation was constructed as follows.

1. The sample was limited to non-military employees (no self employed or employers) with positive usual hours aged 16 to 64 and working within the US (excluding Puerto Rico), and not enrolled in education since February 1, 2000. Individuals with any of the following characteristics allocated by the Census Bureau were also excluded: wage and salary income, hours of work, weeks of work, occupation, commuting time or employment status.

2. Hourly wage rates were constructed by dividing wage and salary income from 1999 by weeks worked in 1999 and usual weekly work hours. Individuals with constructed hourly wage rates above \$200 or below \$1 were excluded.

3. Potential work experience was constructed by subtracting a generated years of schooling variable from age minus 5. This generated years of schooling variable was constructed from reports of highest education level attained. For individuals attaining post-school levels, years were allocated as follows: (1) some college, but less than 1 year = 13, (2) one or more years of college, no degree = 14, (3) associate degree = 15, (4) bachelor's degree = 16, (5) master's degree = 17, (6) professional degree = 18, and (7) doctorate degree = 20. Mid-points were allocated for grade ranges, and for those who attended grade 12 but did not receive a degree were allocated 11.5. For those with less than 9 years of schooling, potential work experience was calculated as age minus 15. The small number of individuals with a negative potential experience calculated were excluded.

4. Part-time workers were indicated by usual weekly work hours of less than 35.

The occupation characteristics constructed using information from O*NET (version 15.0) were constructed as follows.

1. Job zones were in some instances provided only for occupations at a more detailed level than those identified in the US Census. In those instances, the Census occupation group was allocated the highest zone among the individual occupations in O*NET within that group.

2. Hazzards is the simple average of the context (CX) measures (ranging from 1 to 5) for the following exposures in work contexts: contaminants; radiation; disease or infections; high places; hazzardous conditions; hazzardous equipment; and minor burns, cuts, bites or stings.

3. Strength is the simple average of the importance (IM) measures (range 1 to 5) for the following abilities: static strength, explosive strength, dynamic strength and trunk strength.

4. Environment is the simple average of the context (CX) measures (range 1 to 5) for the following work contexts: sounds, noise levels are distracting or uncomfortable; very hot or cold temperatures, extremely bright or inadequate light; cramped work space or awkward positions; and exposed to whole body vibration.

5. The other occupation level characteristics were also measured using the importance (IM) variable. The full titles of these work activities are:

   (a) Computers - Interacting With Computers

   (b) Caring - Assisting and Caring for Others

   (c) Advising - Provide Consultation and Advice to Others

   (d) Teaching - Training and Teaching Others

   (e) Public - Performing for or Working Directly with the Public

   (f) Relationships - Establishing and Maintaining Interpersonal Relationships

The number of fatal injuries by occupation over the period 2003 to 2010 were taken from the BLS's Census of Fatal Occupational Injuries (CFOI). These numbers were transformed into the annual number of fatalities per million work hours within each occupation using information on hours worked from the 2000 US Census micro-data. Usual weekly work hours of employees were multiplied by 52 weeks and summed using Census person weights to construct annual work hours by occupation.

The US SIPP 2004 sample employed in estimation was constructed using the same criteria as outlined for the US Census above, with the following differences.

1. Hourly wages were constructed by using the actual report of the hourly wage rate if one was reported. If no hourly wage rate was reported, or the rate reported was top-coded at $28, then the hourly wage rate was calculated by dividing monthly wage and salary earnings by (usual weekly hours times 365/(12*7)).

2. The SIPP micro-data provides information on up to two jobs held by individuals each wave. Information was employed only for a job that was held at the time of interview. If both reported jobs were held at the time of interview, information from the job with the longest reported weekly hours of work was employed.

3. In around 4% of cases, individuals reported that their weekly hours of work varied, rather than reporting a figure for usual weekly hours. These individuals were controlled for in the estimates using an indicator.

4. There are no PUMA details provided in the SIPP. To control for residence, individual state indicators and these state indicators interacted with an indicator of city residence were included.

5. To control for differences in the time period when individual surveys took place, separate indicators for month times year of interview were included.

References

Adda, Jerome, Christian Dustmann and Katrien Stevens (2010) "The Career Cost of Children," paper presented at the Econometric Society World Congress, Shanghai.

Albrecht, James, Per-Anders Edin, Marianne Sundstrm and Susan Vroman (1999) "Career Interruptions and Subsequent Earnings: A Reexamination Using Swedish Data," *Journal of Human Resources* 34, 294-312.

Antecol, Heather and Deborah A. Cobb-Clark (2010) "Do Non-cognitive Skills Help Explain the Occupational Segregation of Young People?" Melbourne Institute Working Paper no. 13/10.

Blinder, Alan (1973) "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources* 8(4), 436-455.

Bonin, Holger , Thomas Dohmen, Armin Falk, David Huffman and Uwe Sunde (2007) "Cross-sectional earnings risk and occupational sorting: The role of risk attitudes," *Labour Economics* 14(6), 926-937.

Borghans, Lex, Bas ter Weel and Bruce Weinberg (2008) "Interpersonal Styles and Labor Market Outcomes," *Journal of Human Resources* 43(4), 815-858.

Bureau of Labor Statistics (2011) "American Time Use Survey - 2010 Results'" News Release, June 22, Bureau of Labor Statistics, US Department of Labor.

Cavallo, Alex and June O'Neill (2004) "Determinants of the Gender Gap in Occupations and Earnings," Paper presented at the meetings of the Society of Labor Economists, San Antonio, May.

Cobb-Clark, Deborah and Michelle Tan (2011) "Non-cognitive Skills, Occupational Attainment and Relative Wages," *Labour Economics* 18(1), 1-14.

Connolly, Sara and Mary Gregory (2008) "Moving Down: Women's Part-time Work and Occupational Change in Britain 1991-2001," *Economic Journal* 118(526), F52-76.

Connolly, Sara and Mary Gregory (2009) “The part-time pay penalty: earnings trajectories of British women,” *Oxford Economic Papers Supplementary Issue ’Women and Wages’* 61(supplement 1), i76-i97.

Connolly, Sara and Mary Gregory (2010) “Dual tracks: part-time work in life-cycle employment for British women,” *Journal of Population Economics* 23, 907-931.

Dan, Ioana (2010) “Gender based self selection into industries and occupations,” unpublished manuscript, University of Toronto.

Dargay, Joyce M. and Jos Van Ommeren (2005) “The Effect of Income on Commuting Time using Panel Data,” paper presented at the European Regional Science Association (ERSA) 45th annual conference, paper no. ersa05p593.

DeLeire, Thomas and Helen Levy (2004) “Worker Sorting and the Risk of Death on the Job,” *Journal of Labor Economics* 22(4), 925-953.

Fortin, Nicole (2008) “The Gender Wage Gap among Young Adults in the United States: The Importance of Money vs. People,” *Journal of Human Resources* 43(4), 884-918.

Grazier, Suzanne and Peter Sloane (2008) “Accident risk, gender, family status and occupational choice in the UK,” *Labour Economics* 15(5), 938-957.

Golden, Lonnie (2001) “Flexible work schedules: which workers get them?” *American Behavioral Scientist* 44 (7), 1157-1178.

Hersch, Joni (1998) “Compensating Differentials for Gender-Specific Job Injury Risks,” American Economic Review 88(3), 598-607.

Isacsson, Gunnar and Jan-Erik Swardh (2007) “An empirical on-the-job search model with preferences for relative earnings: How high is the value of commuting time?” unpublished manuscript, Swedish National Road and Transport Research Institute.

Jacobsen, Joyce (2007) *The Economics of Gender*, 3rd Edition, Blackwell Publishing, Malden, MA.

Krueger, Alan and David Schkade (2008) "Sorting in the Labor Market: Do Gregarious Workers Flock to Interactive Jobs?" *Journal of Human Resources* 43(4), 859-883.

MacDonald, Heather and Alan H. Peters (1996) "Distance and Labor Force Participation: Implications for Urban and Rural Women," in Sandra Rosenbloom (Ed.) *Women's Travel Issues: Proceedings from the Second National Conference*, Department of Transportation, Washington D.C.

Macpherson, David A. and Barry T. Hirsch (1995) "Wages and Gender Composition: Why Do Women's Jobs Pay Less?" *Journal of Labor Economics* 13(3), 426-471.

Manning, Alan (2003) "The Real Thin Theory: Monopsony in Modern Labour Markets," *Labour Economics* 10(1), 105-131.

McCrate, Elaine (2005) "Flexible hours, workplace authority, and compensating wage differentials in the US," *Feminist Economics* 11 (1), 11-39.

Moretti, Enrico (2011) "Local Labor Markets," in the *Handbook of Labor Economics Volume 4B*, edited by Orley Ashenfelter and David Card, Elsevier BV, 1237-1313.

Oaxaca, Ronald (1973) "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review* 14(3), 693-709.

O'Neill, June E. (2003) "Catching Up: The Gender Gap in Wages, circa 2000," *American Economic Review: AEA Papers and Proceedings* 93(2), 309-314.

Rosenbloom, Joshua, Ronald Ash, Brandon Dupont and LeAnne Coder (2008) "Why are there so few women in information technology? Assessing the role of personality in career choices," *Journal of Economic Psychology* 29, 543-554.

Ross, Stephen L. and Yves Zenou (2004) "Shirking, Commuting and Labor Market Outcomes," Research Institute of Industrial Economics working paper no. 627.

Rupert, Peter, Elena Stancanelli and Etienne Wasmer (2009) "Commuting, Wages and Bargaining Power," IZA Discussion Paper no. 4510.

So, Kim S., Peter F. Orazem and Daniel M. Otto (2001) "The Effects of Housing Prices, Wages, and Commuting Time on Joint Residential and Job Location Choices," *American Journal of Agricultural Economics* 83(4), 1036-1048.

Stutzer, Alois and Bruno S. Frey (2008) "Stress that Doesn't Pay: The Commuting Paradox," *Scandinavian Journal of Economics* 110(2), 339-366.

Timothy, Darren and William C. Wheaton (2001) "Intra-Urban Wage Variation, Employment Location, and Commuting Times," Journal of Urban Economics 50(2), 338-366.

van Ommeren, Jos and Mogens Fosgerau (2009) "Workers' marginal costs of commuting," *Journal of Urban Economics* 65(1), 38-47.

Table 1: **Summary statistics - 2000 US Census**

	Females		Males	
	mean	st. dev.	mean	st. dev.
Hourly wage	15.56	12.65	20.51	18.14
Years of schooling*	13.90	2.38	13.67	2.75
Age	40.21	10.98	39.70	11.03
Part-time	0.190		0.049	
Married	0.604		0.666	
Immigrant	0.112		0.142	
Black	0.100		0.076	
Hispanic	0.081		0.106	
American native	0.013		0.012	
Asian	0.040		0.041	
Pacific Islander	0.002		0.002	
Other race	0.041		0.057	
Non-profit employee	0.119		0.048	
Federal govt. employee	0.031		0.036	
State govt. employee	0.066		0.044	
Local govt. employee	0.100		0.067	
Commute (minutes)	23.68	21.36	27.70	25.46
work at home	0.013		0.011	
leave for work				
- 0 to 5:59 am	0.072		0.157	
- 6 to 6:59 am	0.168		0.251	
- 7 to 7:59 am	0.364		0.291	
- 8 to 8:59 am	0.198		0.120	
- 9 to 11:59 am	0.086		0.057	
- 12 to 6:59 pm	0.076		0.081	
- 7 to 11:59 pm	0.022		0.033	
Observations	1,494,545		1,621,748	

Notes: * Constructed from highest education attained categories (see Appendix A). Averages constructed using population weights provided with the Census data.

Table 2: **Individual level regressions (step 1)**

	Females		Males	
	coeff.	t-stat	coeff.	t-stat
4th grade or less	-0.209	-20.1	-0.246	-33.4
5th or 6th grade	-0.182	-26.4	-0.197	-42.7
7th or 8th grade	-0.147	-24.2	-0.157	-37.5
9th grade	-0.139	-25.8	-0.165	-43.7
10th grade	-0.113	-27.3	-0.126	-39.5
11th grade	-0.093	-22.5	-0.101	-32.7
12th grade, no diploma	-0.043	-12.5	-0.046	-17.0
Some college (<1 year)	0.055	33.7	0.054	32.1
College ≥ 1 yr, no degree	0.083	59.2	0.082	60.1
Associate degree	0.120	70.9	0.114	64.2
Bachelor's degree	0.251	158.4	0.241	147.1
Master's degree	0.436	193.0	0.367	151.4
Professional degree	0.321	61.9	0.386	61.3
Doctorate degree	0.508	81.7	0.471	93.7
Potential experience (PE)	0.057	67.6	0.060	67.2
PE^2 / 10	-0.028	-40.0	-0.024	-33.1
PE^3 / 1,000	0.062	28.3	0.049	21.4
PE^4 / 100,000	-0.052	-22.2	-0.041	-17.1
Part-time	-0.007	-5.06	0.050	15.6
Married	0.008	9.03	0.121	118.6
Immigrant	-0.076	-38.7	-0.091	-47.3
Black	-0.012	-6.31	-0.064	-32.2
Hispanic	-0.020	-7.9	-0.048	-20.3
American native	-0.055	-13.2	-0.065	-15.1
Asian	-0.011	-3.75	-0.049	-16.1
Pacific Islander	-0.015	-1.53	-0.019	-1.83
Other race	-0.021	-6.88	-0.023	-8.72

Notes: Table continued on next page.

Table 2 (cont.): **Individual level regressions (step 1)**

	Females		Males	
	coeff.	t-stat	coeff.	t-stat
Non-profit employee	-0.010	-5.78	-0.067	-24.2
Federal govt. employee	0.159	51.1	0.094	31.3
State govt. employee	0.074	32.6	-0.015	-5.69
Local govt. employee	0.056	26.0	-0.006	-2.71
Commute (minutes)	0.0035	32.0	0.0022	20.6
Commute^2 / 100	-0.0029	-13.3	-0.0019	-9.64
Commute^3 / 10,000	0.0007	7.02	0.0006	6.29
work at home	0.037	6.53	0.044	7.19
leave for work				
- 0 to 5:59 am	0.027	13.4	0.020	10.2
- 6 to 6:59 am	0.026	17.5	0.002	1.29
- 7 to 7:59 am	0.017	13.8	0.000	-0.12
- 9 to 11:59 am	-0.032	-15.4	-0.050	-19.0
- 12 to 6:59 pm	0.006	2.69	-0.010	-4.25
- 7 to 11:59 pm	0.048	14.6	0.019	6.85
Occupation (470)	Yes		Yes	
Industry (24)	Yes		Yes	
PUMA (2071)	Yes		Yes	
Observations	1,494,545		1,621,748	
R-squared	0.4166		0.4446	

Notes: Dependent variable is the natural log of the hourly wage rate. Estimates constructed using Census provided population weights. T-statistics based on White robust standard errors.

Table 3: **Occupation variable summary statistics**

	Females		Males	
	mean	st. dev.	mean	st. dev.
Proportion female	0.666	0.242	0.304	0.239
Prop. Part-Time (males)	0.085	0.072	0.049	0.052
Mean hours FT (males)	44.31	2.35	45.42	2.85
Mean commute (males)	26.44	3.48	27.70	4.45
Job zone 1	0.057	0.232	0.052	0.222
Job zone 2	0.294	0.456	0.323	0.468
Job zone 3	0.285	0.452	0.292	0.455
Job zone 4	0.226	0.419	0.236	0.425
Job zone 5	0.138	0.346	0.097	0.296
Hazzards	1.658	0.472	2.006	0.648
Strength required	1.615	0.522	1.863	0.595
Environmental bads	1.817	0.375	2.309	0.692
Fatal injuries (mill. hours)	0.010	0.027	0.038	0.073
Computers	3.619	0.909	3.249	1.002
Caring	3.091	0.788	2.705	0.598
Advising	2.654	0.628	2.754	0.619
Teaching	3.144	0.614	3.123	0.526
Public	3.234	0.828	2.892	0.910
Relationships	3.821	0.444	3.622	0.526
Observations	470		470	

Notes: Sources: US Census 2000, O*NET and US Bureau of Labor Statistics. Means were constructed using the US Census provided individual person weights.

Table 4: **Occupation level regressions - Females**

	Model one		Model two		Model three		Model four	
	coeff.	t-stat	coeff.	t-stat	coeff.	t-stat	coeff.	t-stat
Proportion female (PF)	-0.159	-5.81	-0.039	-1.29	-0.255	-6.61	-0.191	-4.75
Prop. Part-Time (male)			-0.359	-3.14			-0.259	-2.32
Mean hours FT (male)			0.014	4.63			0.011	3.22
Mean commute (male)			0.014	7.30			0.013	6.74
Job zone 2	0.141	4.38	0.078	2.46	0.038	1.24	0.022	0.72
Job zone 3	0.288	8.93	0.192	5.88	0.099	2.95	0.067	2.05
Job zone 4	0.393	12.0	0.268	7.81	0.166	4.44	0.142	3.93
Job zone 5	0.474	13.6	0.358	9.89	0.212	5.27	0.183	4.71
Hazzards					0.110	4.41	0.102	4.33
Strength required					-0.063	-2.57	-0.040	-1.70
Environmental bads					-0.049	-1.85	-0.069	-2.72
Fatal injuries (mill. hrs)					-0.039	-0.36	-0.153	-1.46
Computers					0.064	6.20	0.049	4.59
Caring					0.042	2.77	0.051	3.56
Advising					0.055	3.09	0.037	2.18
Teaching					-0.060	-3.96	-0.037	-2.50
Public					0.009	0.99	0.021	2.46
Relationships					0.047	2.10	0.012	0.55
Observations	470		470		470		470	
R-squared	0.4763		0.5637		0.6100		0.6566	

Notes: First step estimates of sample variance of occupation indicator coefficients used as weights.

Table 5: **Occupation level regressions - Males**

	Model one		Model two		Model three		Model four	
	coeff.	t-stat	coeff.	t-stat	coeff.	t-stat	coeff.	t-stat
Proportion female (PF)	-0.133	-5.46	0.000	-0.01	-0.178	-4.91	-0.103	-3.02
Prop. Part-Time (male)			-0.848	-6.66			-0.682	-5.12
Mean hours FT (male)			0.013	5.83			0.010	4.42
Mean commute (male)			0.009	6.98			0.010	8.04
Job zone 2	0.133	4.29	0.041	1.38	0.067	2.21	0.008	0.28
Job zone 3	0.253	8.11	0.130	4.22	0.115	3.48	0.045	1.46
Job zone 4	0.406	12.9	0.247	7.66	0.188	5.03	0.123	3.54
Job zone 5	0.448	12.8	0.296	8.45	0.223	5.56	0.154	4.14
Hazzards					0.052	2.24	0.060	2.83
Strength required					-0.069	-3.12	-0.035	-1.75
Environmental bads					-0.002	-0.07	-0.043	-2.18
Fatal injuries (mill. hrs)					-0.015	-0.18	-0.167	-2.15
Computers					0.031	3.16	0.028	3.06
Caring					0.008	0.59	0.014	1.11
Advising					0.051	2.93	0.034	2.15
Teaching					-0.012	-0.75	0.000	-0.01
Public					-0.003	-0.40	0.016	2.14
Relationships					0.055	2.76	0.008	0.45
Observations	470		470		470		470	
R-squared	0.4968		0.6117		0.5856		0.6677	

Notes: First step estimates of sample variance of occupation indicator coefficients used as weights.

Table 6: **Correlation among occupational measures**

	PF	PT	FT hrs	Commute	Hazz.	Strgth	Envir.
Prop. Female (PF)	1						
Proportion PT	0.45	1					
FT hours	-0.34	-0.37	1				
Commute	-0.26	-0.33	0.02	1			
Hazzards	-0.49	-0.06	-0.12	-0.01	1		
Strength	-0.36	0.19	-0.19	-0.18	0.81	1	
Environment	-0.67	-0.14	-0.06	0.14	0.87	0.74	1
Fatal injuries	-0.41	0.01	0.10	0.23	0.38	0.38	0.51
Computers	0.32	-0.31	0.09	0.18	-0.64	-0.77	-0.60
Caring	0.44	0.30	-0.06	-0.33	0.15	0.25	-0.10
Advising	-0.13	-0.40	0.46	0.09	-0.16	-0.31	-0.21
Teaching	0.03	-0.09	0.20	-0.27	0.10	0.07	0.01
Public	0.32	0.31	0.13	-0.28	-0.14	0.04	-0.16
Relationships	0.33	-0.18	0.36	0.04	-0.52	-0.57	-0.57

	Fatal	Comput.	Care	Advise	Teach	Public	Relations
Fatal injuries	1						
Computers	-0.38	1					
Caring	-0.03	-0.07	1				
Advising	-0.12	0.42	0.19	1			
Teaching	-0.04	0.11	0.44	0.55	1		
Public	0.01	0.02	0.52	0.06	0.13	1	
Relationships	-0.26	0.55	0.33	0.60	0.28	0.35	1

Notes: Sources: US Census 2000, O*NET and US Bureau of Labor Statistics. Correlations were constructed using the US Census provided individual person weights.

Table 7: **Decomposition of mean log wage gap - individual coefficients**

	log points	per cent
Raw log wage gap	0.242	100.0
Education	-0.009	-3.7
Potential experience	0.000	-0.1
Part-time	-0.003	-1.2
Married	0.004	1.7
Immigrant	-0.003	-1.1
Race	0.000	-0.1
Employer	0.002	0.8
Commute	0.005	2.0
Work at home	0.000	0.0
Start time	0.006	2.5
Industry	0.030	12.4
TOTAL	0.032	13.1

Notes: Decomposition employs average of male and female regression coefficients.

Table 8: **Decomposition of mean log wage gap - occupation coefficients**

	Model 1		Model 2		Model 3		Model 4	
	level	%	level	%	level	%	level	%
Proportion female (PF)	0.053	21.8	0.007	2.9	0.078	32.4	0.053	22.0
Prop. Part-Time (male)			0.021	8.9			0.017	6.9
Mean hours FT (male)			0.015	6.2			0.012	4.8
Mean commute (male)			0.014	5.9			0.015	6.1
Job zone	-0.009	-3.8	-0.008	-3.3	-0.005	-2.0	-0.005	-2.0
Hazzards, strength, env.					-0.001	-0.3	-0.013	-5.5
Fatal injuries					-0.001	-0.3	-0.005	-1.9
Computers					-0.018	-7.3	-0.014	-5.9
Care, teach, relate etc					-0.013	-5.4	-0.027	-11.1

Notes: Decomposition employs average of male and female regression coefficients.

Table 9: **SIPP Wage Regressions - Females**

	Pooled		Fixed Effects	
	coeff.	t-stat	coeff.	t-stat
Less than 1st grade	-0.282	-2.53	-0.064	-0.95
1st to 4th grade	-0.189	-3.82	-0.050	-1.49
5th or 6th grade	-0.175	-6.20	0.020	1.19
7th or 8th grade	-0.174	-5.99	-0.022	-0.78
9th grade	-0.094	-3.34	0.037	1.29
10th grade	-0.117	-4.38	0.011	0.59
11th grade	-0.103	-6.07	-0.057	-3.33
12th grade, no diploma	-0.063	-2.76	0.011	0.36
Some college, no degree	0.056	6.33	0.001	0.06
Certificate / diploma	0.022	1.84	0.006	0.28
Associate degree	0.160	8.55	0.043	1.60
Bachelor's degree	0.298	16.41	0.064	2.40
Master's degree	0.462	19.90	0.073	1.98
Professional degree	0.571	12.15	0.073	0.84
Doctorate degree	0.595	12.03	0.114	1.44
Potential experience	0.042	6.58		
PE^2 / 10	-0.019	-3.93	0.009	3.65
PE^3 / 1,000	0.040	2.66	-0.039	-3.76
PE^4 / 100,000	-0.033	-2.02	0.041	3.24
Part-time	-0.093	-5.33	0.033	6.07
Hours vary	-0.130	-8.51	0.002	0.48
Married	0.031	4.89	0.010	1.51
Black	-0.041	-3.79		
Asian	-0.072	-3.92		
Hispanic	-0.104	-8.37		
Non-profit employee	-0.008	-0.44	0.011	0.94
Federal govt. employee	0.207	8.71	0.077	2.72
State govt. employee	-0.001	-0.05	0.042	2.16
Local govt. employee	0.031	1.60	0.012	0.67

Table 9 (cont.): **SIPP Wage Regressions - Females**

	Pooled		Fixed Effects	
	coeff.	t-stat	coeff.	t-stat
Proportion female	-0.054	-0.69	0.035	1.32
Prop. Part-Time (male)	-0.432	-2.19	-0.603	-8.88
Mean hours FT (male)	0.013	2.38	0.002	1.00
Mean commute (male)	0.018	6.64	0.004	2.85
Job zone 2	0.084	1.43	0.068	4.76
Job zone 3	0.107	1.91	0.082	5.10
Job zone 4	0.225	3.92	0.098	5.29
Job zone 5	0.294	4.07	0.103	4.93
Hazzards	0.177	3.51	0.048	2.91
Strength required	-0.106	-2.18	-0.085	-6.45
Environmental bads	-0.007	-0.15	0.025	1.53
Fatal injuries (mill. hrs)	-0.011	-0.05	-0.138	-1.05
Computers	0.037	1.51	-0.017	-2.77
Caring	0.037	1.62	0.010	1.20
Advising	0.041	1.59	0.017	1.73
Teaching	-0.031	-1.53	-0.012	-1.39
Public	0.005	0.35	-0.001	-0.27
Relationships	0.024	0.72	-0.005	-0.42
Industry (20)	Yes		Yes	
State (51)	Yes		Yes	
City times state (47)	Yes		Yes	
Month of survey (48)	Yes		Yes	
Observations	146,019		146,019	
Individuals			24,463	
R-squared within			0.0737	
R-squared between			0.2840	
R-squared TOTAL	0.4956		0.2523	

Notes: The pooled regression t-statistics allow for clustering at the occupation level. The fixed effects t-statistics are based on White robust standard errors.

Table 10: **SIPP Wage Regressions - Males**

	Pooled		Fixed Effects	
	coeff.	t-stat	coeff.	t-stat
Less than 1st grade	-0.311	-4.44	0.141	2.67
1st to 4th grade	-0.212	-8.01	0.020	0.59
5th or 6th grade	-0.173	-7.53	-0.017	-0.45
7th or 8th grade	-0.151	-7.11	0.007	0.34
9th grade	-0.142	-5.81	0.019	0.58
10th grade	-0.069	-4.44	-0.002	-0.08
11th grade	-0.048	-2.86	-0.014	-0.84
12th grade, no diploma	-0.086	-4.08	-0.081	-3.75
Some college, no degree	0.091	7.93	0.029	2.16
Certificate / diploma	0.069	5.15	0.043	1.51
Associate degree	0.139	10.06	0.040	1.51
Bachelor's degree	0.274	15.45	0.049	1.66
Master's degree	0.397	16.78	0.140	3.15
Professional degree	0.554	13.00	0.067	0.74
Doctorate degree	0.526	15.29	0.195	1.76
Potential experience	0.038	6.67		
PE^2 / 10	-0.010	-1.98	0.013	5.47
PE^3 / 1,000	0.006	0.35	-0.066	-6.01
PE^4 / 100,000	0.003	0.15	0.080	5.51
Part-time	-0.170	-10.75	0.021	2.30
Hours vary	-0.089	-8.22	-0.018	-3.23
Married	0.110	16.48	0.013	1.68
Black	-0.111	-9.68		
Asian	-0.087	-5.16		
Hispanic	-0.156	-12.83		
Non-profit employee	-0.100	-3.15	-0.039	-1.64
Federal govt. employee	0.201	8.68	0.037	0.98
State govt. employee	-0.047	-2.27	0.021	0.67
Local govt. employee	-0.011	-0.53	0.029	1.21

Table 10 (cont.): **SIPP Wage Regressions - Males**

	Pooled		Fixed Effects	
	coeff.	t-stat	coeff.	t-stat
Proportion female	-0.124	-1.94	-0.003	-0.11
Prop. Part-Time (male)	-0.685	-4.16	-0.628	-6.31
Mean hours FT (male)	0.011	2.71	0.000	-0.12
Mean commute (male)	0.012	5.52	0.003	3.26
Job zone 2	0.017	0.76	0.020	1.44
Job zone 3	0.067	2.65	0.038	2.46
Job zone 4	0.131	3.81	0.049	2.33
Job zone 5	0.191	3.86	0.075	3.12
Hazzards	0.072	2.84	0.022	1.47
Strength required	-0.075	-2.57	-0.048	-3.43
Environmental bads	-0.040	-1.69	-0.017	-1.19
Fatal injuries (mill. hrs)	-0.146	-1.35	0.009	0.15
Computers	0.019	1.59	-0.003	-0.55
Caring	0.013	0.62	-0.001	-0.11
Advising	0.054	2.58	-0.001	-0.08
Teaching	-0.025	-1.10	0.012	1.15
Public	-0.013	-1.25	-0.014	-2.56
Relationships	0.049	1.90	0.007	0.58
Industry (20)	Yes		Yes	
State (51)	Yes		Yes	
City times state (47)	Yes		Yes	
Month of survey (48)	Yes		Yes	
Observations	146,398		146,398	
Individuals			24,244	
R-squared within			0.0588	
R-squared between			0.2106	
R-squared TOTAL	0.4925		0.1924	

Notes: The pooled regression t-statistics allow for clustering at the occupation level. The fixed effects t-statistics are based on White robust standard errors.

Table 11: **Occupation variable summary statistics - extended model**

	Females		Males	
	mean	st. dev.	mean	st. dev.
FT hours coef. variation (m)	0.170	0.022	0.170	0.024
Commute coef. variation (m)	0.887	0.083	0.899	0.088
Prop. work home (males)	0.012	0.017	0.011	0.016
Proportion leave (males)				
- 0 to 5:59 am	0.098	0.070	0.157	0.100
- 6 to 6:59 am	0.202	0.076	0.251	0.096
- 7 to 7:59 am	0.328	0.138	0.291	0.126
- 8 to 8:59 am	0.153	0.086	0.120	0.089
- 9 to 11:59 am	0.078	0.068	0.057	0.057
- 12 to 6:59 pm	0.095	0.097	0.081	0.082
- 7 to 11:59 pm	0.032	0.038	0.033	0.040
Concentration in leaving time (m)	0.251	0.069	0.253	0.056
Observations	470		470	

Notes: Sources: US Census 2000. Means were constructed using the US Census provided individual person weights.

Table 12: **Extended occupation level regressions**

	Females		Males	
	coeff.	t-stat	coeff.	t-stat
Proportion female (PF)	-0.175	-4.67	-0.110	-3.24
Prop. Part-Time (male)	0.029	0.23	-0.568	-3.48
Mean hours FT (male)	0.027	6.28	0.019	4.74
Mean commute (male)	0.015	7.76	0.009	6.41
Job zone 2	0.004	0.15	0.017	0.59
Job zone 3	0.045	1.42	0.057	1.84
Job zone 4	0.098	2.75	0.123	3.49
Job zone 5	0.143	3.67	0.142	3.71
Hazzards	0.091	4.12	0.050	2.41
Strength required	-0.014	-0.63	-0.042	-2.14
Environmental bads	-0.060	-2.53	-0.018	-0.90
Fatal injuries (mill. hrs)	-0.082	-0.84	-0.083	-1.07
Computers	0.050	4.80	0.015	1.60
Caring	0.054	3.97	0.019	1.53
Advising	0.025	1.61	0.028	1.83
Teaching	-0.032	-2.27	0.002	0.12
Public	0.026	2.83	0.018	2.10
Relationships	0.001	0.04	0.010	0.53
FT hours coef. variation (male)	-0.990	-2.61	-0.541	-1.29
Commute coef. variation (male)	-0.396	-4.57	-0.389	-4.45
Prop. work home (male)	1.256	2.92	0.433	0.82
Leave 0 to 6:59 am (male)	-0.058	-0.35	-0.301	-1.72
Leave 6 to 6:59 am (male)	0.210	1.62	-0.129	-0.92
Leave 7 to 7:59 am (male)	0.046	0.28	-0.202	-1.13
Leave 9 to 11:59 am (male)	0.263	1.24	-0.095	-0.33
Leave 12 to 6:59 pm (male)	0.094	0.62	0.156	0.99
Leave 7 to 11:59 pm (male)	0.913	3.60	0.085	0.32
Concentration leaving time (male)	0.896	6.31	0.475	3.47
Observations	470		470	
R-squared	0.7256		0.6824	

Notes: First step estimates of sampling variation of occupation indicator coefficients used as weights.

Table A1: **Occupations with the shortest and longest commutes**

Occupation	time
Ten shortest commutes	
Clergy	12.34
Farm, ranch, & other agricultural managers	16.44
Religious workers, other	17.18
Funeral Directors	17.23
Directors, religious activities & education	17.29
Animal trainers	17.36
Funeral service workers	19.32
Bartenders	19.64
Barbers	19.66
Preschool & kindergarten teachers	19.74
Ten longest commutes	
Financial examiners	45.72
Hoist & winch operators	46.32
Rail-track laying & maintenance equipment operators	46.34
Elevator installers & repairers	46.84
Transportation attendants	47.23
Miscellaneous extraction workers	47.91
Sailors & marine oilers	55.87
Ship & boat captains and operators	59.17
Aircraft pilots & flight engineers	64.03
Derrick, rotary drill, service unit operators, roustabouts - oil, gas, & mining	75.87

Notes: 2000 US Census, commutes based on male employees only.

Figure 1: **Occupation Dispersion and Average Commuting Times**

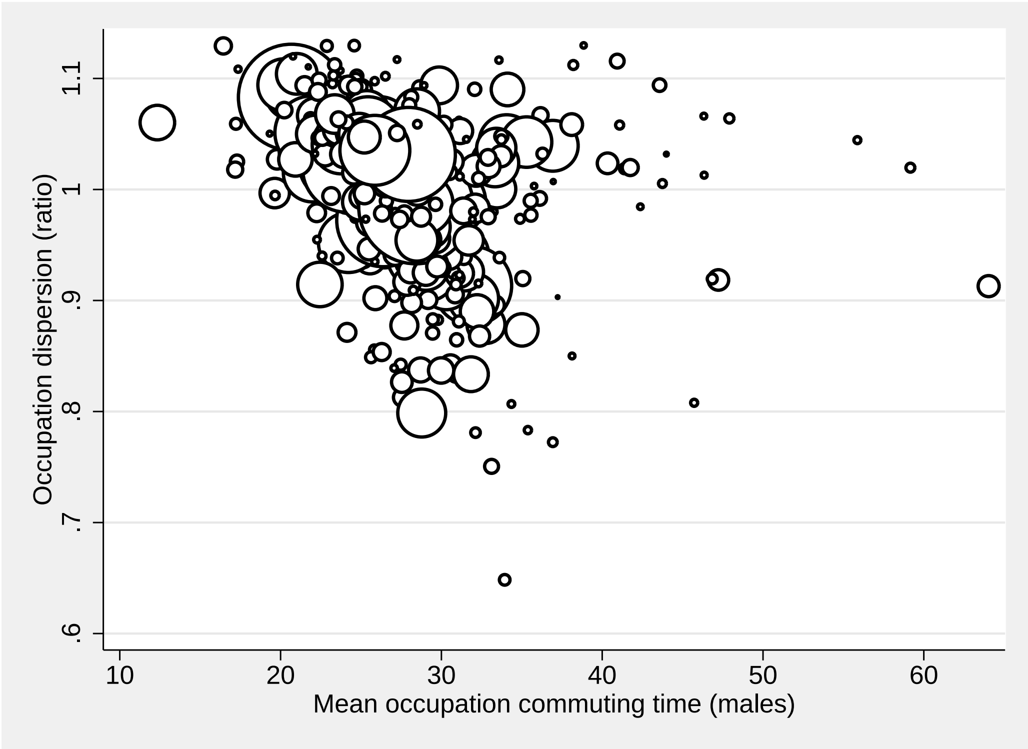


Figure 2: **Predicted Log Wages and Individual Commuting Times**

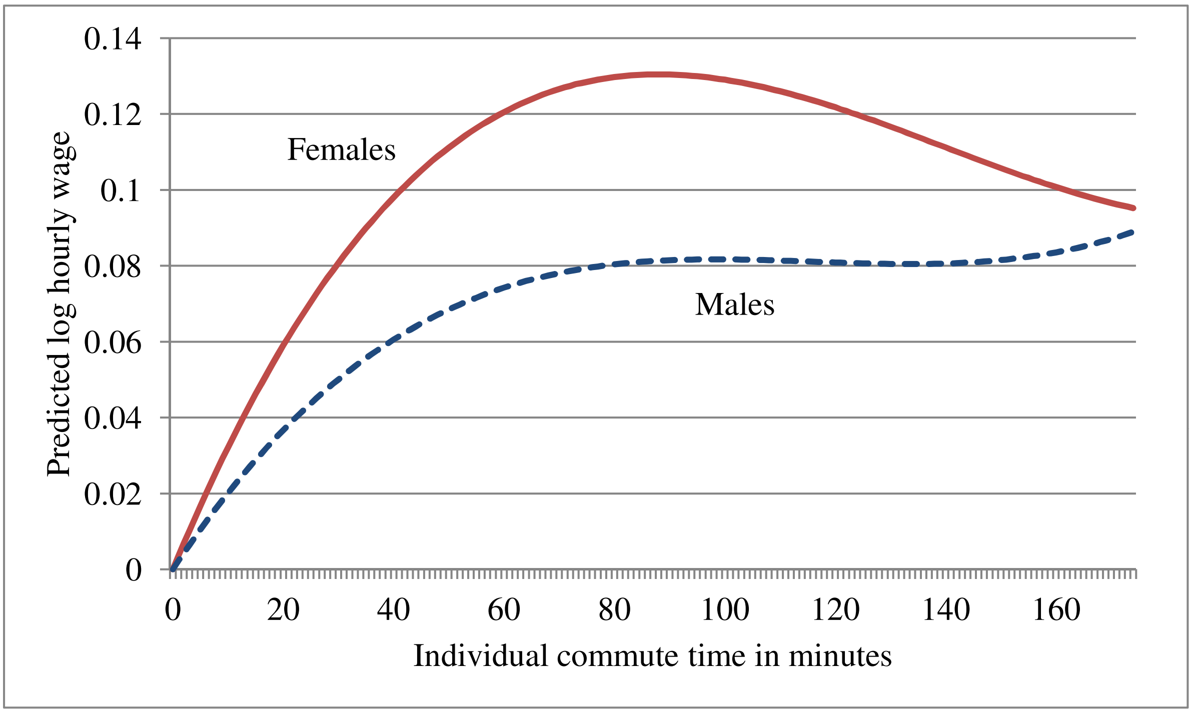
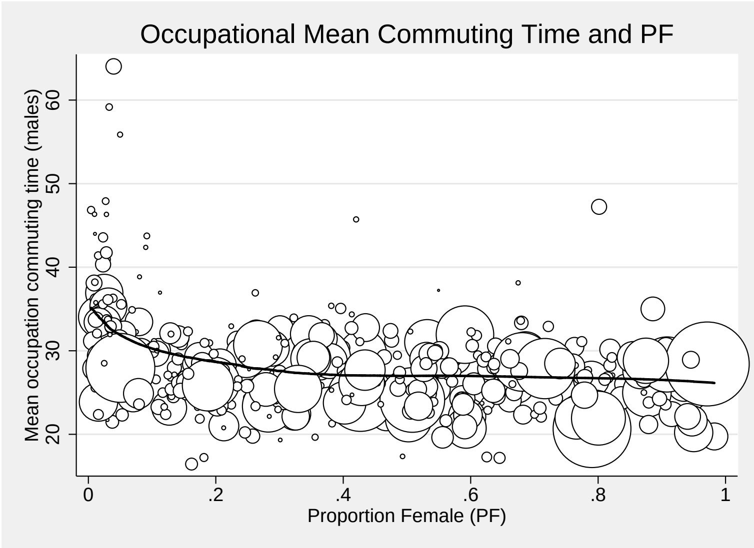
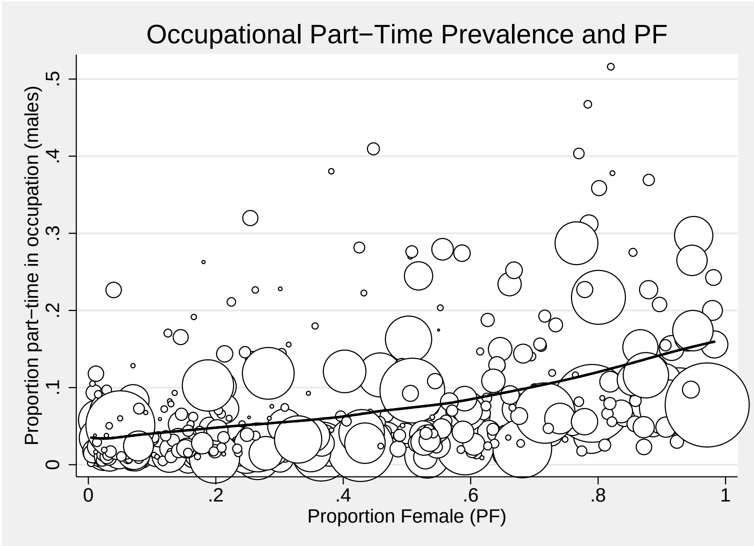


Figure 3: **Family friendly occupation characteristics and proportion female (PF)**





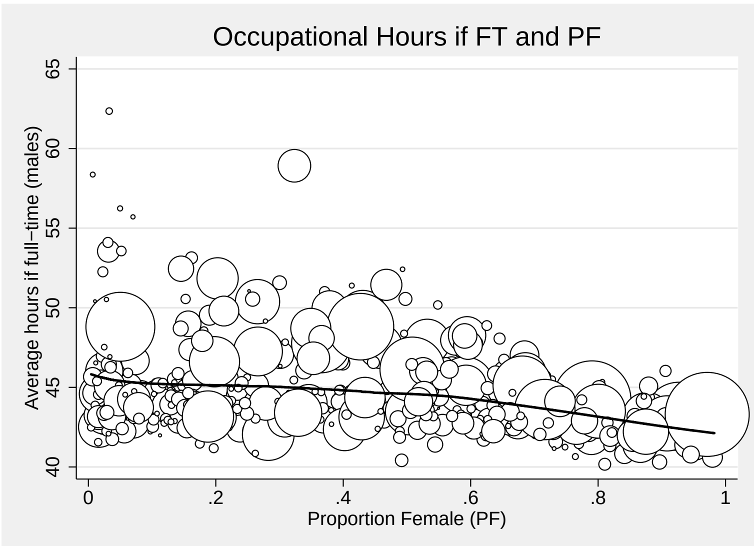


Figure 4: **City-based measures of commuting time by occupation**

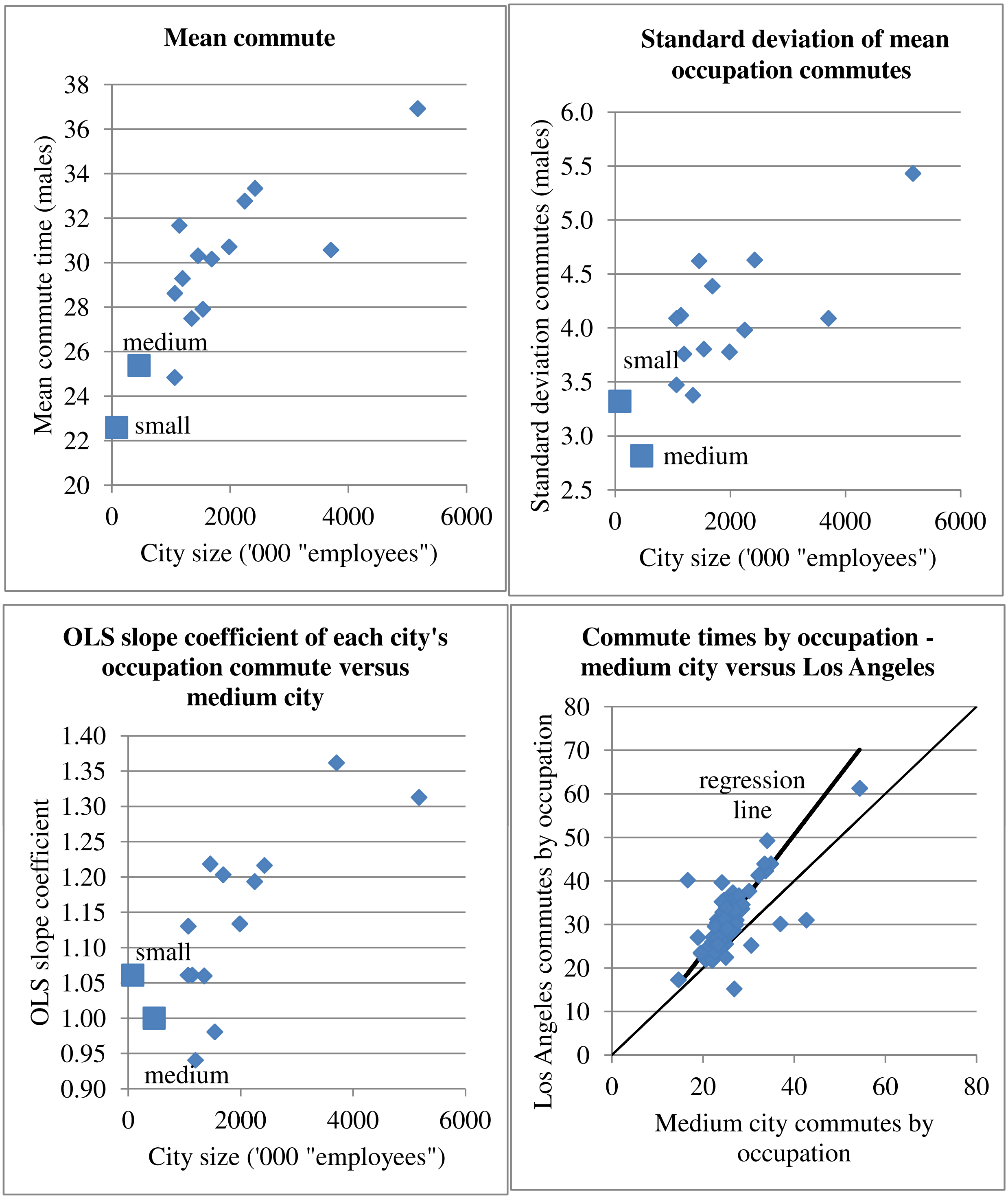


Figure 5a: **Part-time prevalence in the occupation**

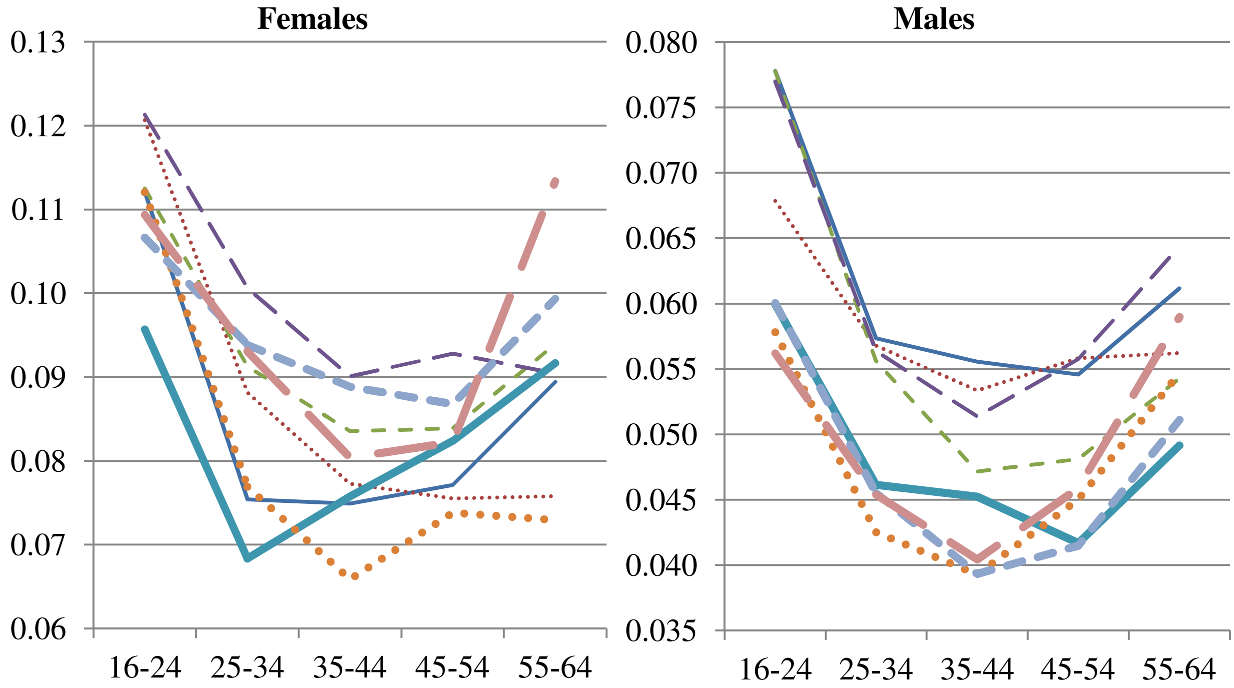
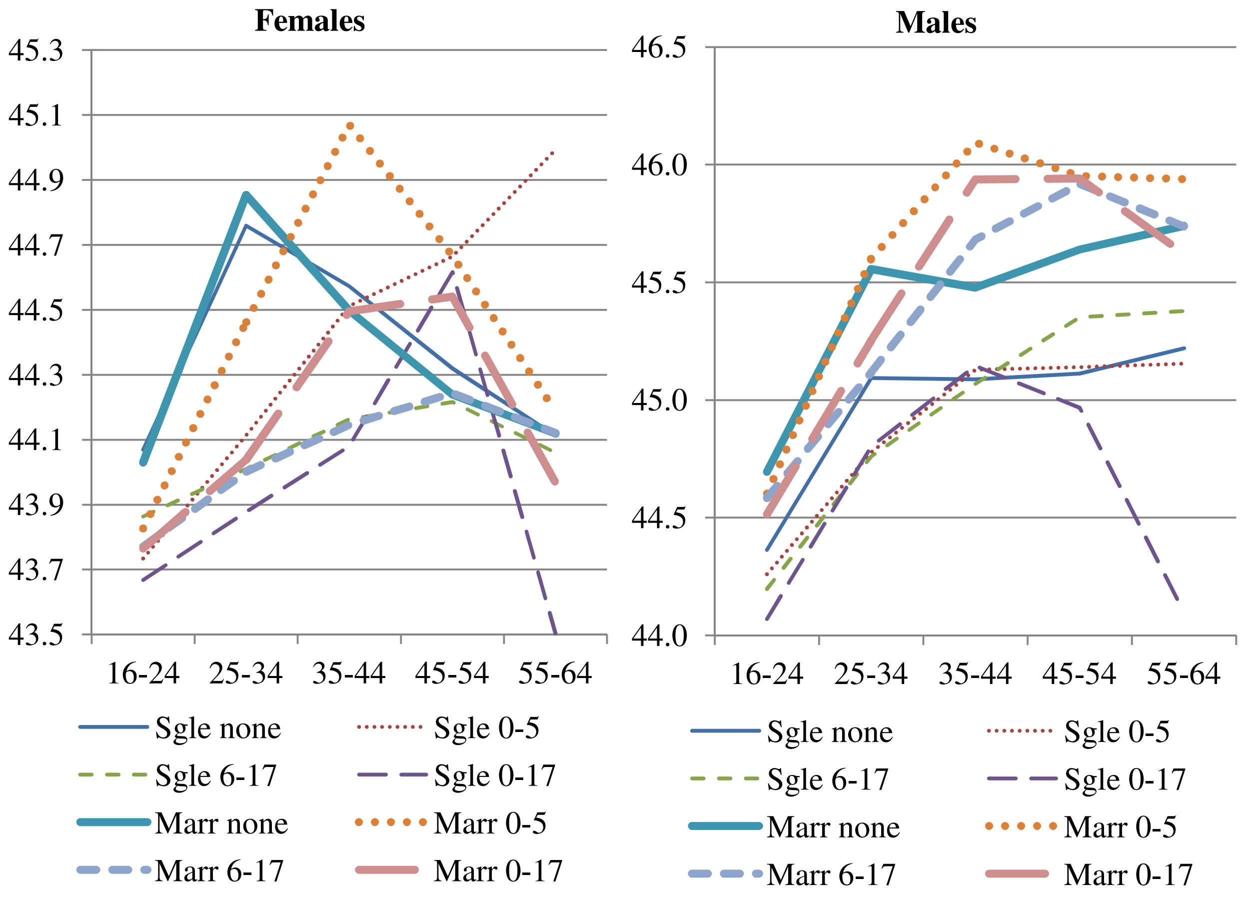
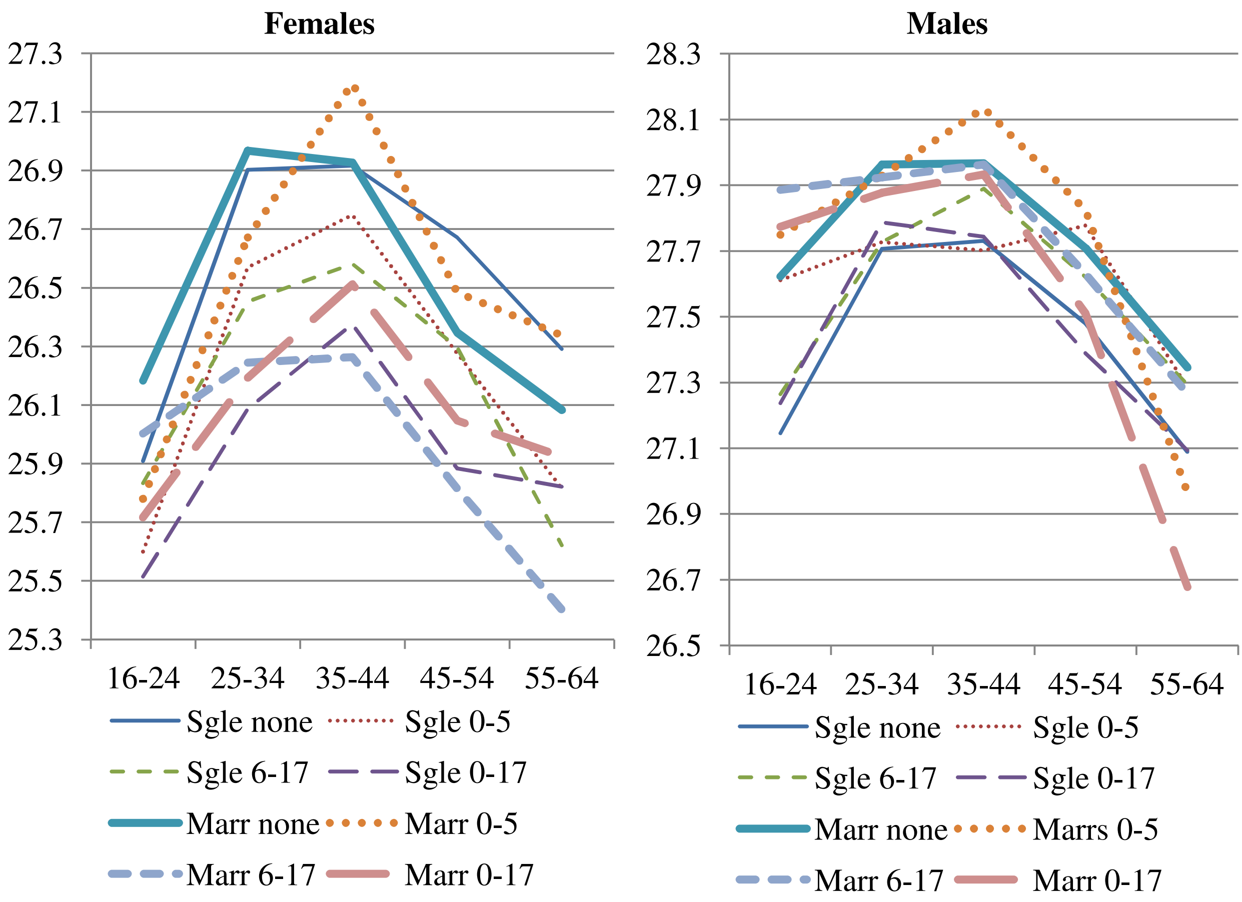


Figure 5b: **Average hours if full-time in the occupation**



Notes: Each data point is the average of the occupation characteristic for the group.
Sgle = not married, Marr = married,
0-5 = children aged 0 to 5 only in the home none = no children under age 18 in the home,
0-17 = children aged both 0 to 5 and 6 to 17 in the home

Figure 5c: **Average commuting time in the occupation**



Notes: Each data point is the average of the occupation characteristic for the group.
Sgle = not married, Marr = married,
0-5 = children aged 0 to 5 only in the home, 6-17 = children aged 6 to 17 only in the home,
0-17 = children aged both 0 to 5 and 6 to 17 in the home