

IFN Working Paper No. 753, 2008

Are Shirking and Leisure Substitutable? An Empirical Test of Efficiency Wages Based on Urban Economic Theory

Stephen L. Ross and Yves Zenou

Are Shirking and Leisure Substitutable? An Empirical Test of Efficiency Wages based on Urban Economic Theory*

Stephen L. Ross[†] Yves Zenou[‡]

June 12, 2008

Abstract

Recent theoretical work has examined the spatial distribution of unemployment using the efficiency wage model as the mechanism by which unemployment arises in the urban economy. This paper extends the standard efficiency wage model in order to allow for behavioral substitution between leisure time at home and effort at work. In equilibrium, residing at a location with a long commute affects the time available for leisure at home and therefore affects the trade-off between effort at work and risk of unemployment. This model implies an empirical relationship between expected commutes and labor market outcomes, which is tested using the Public Use Microdata sample of the 2000 U.S. Decennial Census. The empirical results suggest that efficiency wages operate primarily for blue collar workers, i.e. workers who tend to be in occupations that face higher levels of supervision. For this subset of workers, longer commutes imply higher levels of unemployment and higher wages, which are both consistent with shirking and leisure being substitutable.

JEL Classification: J41, R14.

Keywords: Efficiency wage, leisure, urban unemployment.

*We would like to thank the editor, Richard Arnott, and two anonymous referees for very helpful comments. We are also grateful to Jeffrey Zax and Deborah Garvey as well as the participants of the 2004 SOLE conference for excellent comments. Yves Zenou thanks the Bank of Sweden Tercentenary Foundation for financial support.

[†]Department of Economics, University of Connecticut, 341 Mansfield Road U-63, Storrs, CT 06269, USA.
E-mail: stephen.l.ross@uconn.edu

[‡]Stockholm University and Research Institute of Industrial Economics. E-mail: yves.zenou@ne.su.se

1 Introduction

Many U.S. metropolitan areas as well as European cities are characterized by a concentration of poverty and unemployment in specific regions of their central cities and inner ring suburbs. The concentration of poverty and unemployment in a neighborhood may have external effects on other neighborhood residents leading to poor outcomes in education and family structure, and further exacerbating negative labor market outcomes.¹ Recently, a body of theoretical work has developed that explores the spatial distribution of urban unemployment. A common approach in the literature is to use the efficiency wage model as the mechanism by which unemployment arises in the urban economy. For example, Zenou and Smith (1995) develop a model in which housing prices and workers' location (land market), as well as wages and unemployment (labor market) are determined in equilibrium, and Brueckner and Zenou (2003) examine the impact of job decentralization or spatial mismatch on unemployment using a similar efficiency wage model.² However, up to this point, no empirical work has been conducted to test whether efficiency wage models can explain the spatial distribution of unemployment and earnings within urban labor markets

This paper extends the standard efficiency wage model in order to allow for behavioral substitution between leisure time at home and effort or shirking at work. Given that shirking is a form of leisure, any change to an individual's time endowment is likely to affect the utility from shirking. Further, unless time endowment differences are completely observed by firms, firms will not pay wages that are below the level required to eliminate all shirking in equilibrium in contrast to the standard efficiency wage model (Shapiro and Stiglitz, 1984). This new efficiency wage model with substitution is imbedded into a standard urban model with commuting and a land market. In equilibrium, residing at a location with a long commute affects the trade-off between effort at work and the frequency of unemployment spells by reducing the time available for leisure. This model suggests that either workers effort provided at work or wages vary based upon a worker's residential location, depending upon whether firms can discriminate based on residential location. Such an extension to allow for substitution between leisure and shirking would appear to be especially relevant for urban models of unemployment because these models have been developed to understand the spatial distribution of unemployment and earnings and such models imply substantial differences in

¹A considerable body of research has developed documenting the impact of residential location on employment outcomes; a few recent examples include Bayer, Ross, and Topa (2004) and Topa (2001) on social interactions, Weinberg, Reagan and Yankow (2005) and Katz, Kling, and Liebman (2001) on neighborhood quality, and Weinberg (2000) and Ross (1998) on job access.

²See Zenou (2008) for an overview on urban unemployment and efficiency wages.

time endowments net of commuting across residential locations. Further, this model implies the existence of a relationship between expected commutes and either employment and/or wages, which we will test empirically in this paper.

In the empirical portion of the paper, we estimate the relationship between commutes and employment or earnings using an instrumental variable model. Specifically, we instrument for an individual worker's commute based on across metropolitan differences in the likelihood of each worker choosing specific housing features. The models are estimated using workers in large and mid-sized metropolitan areas drawn from the 2000 Public Use Microdata Sample of the U.S. Decennial Census. The workers are divided into heavily supervised and lightly supervised occupations, and separate models are estimated for the two samples. For workers in heavily supervised occupations, we find that unemployment and wages both increase with commute times, which is consistent with substitutability between shirking and leisure and with firms being partially able to base wages on commutes in order to reduce shirking among workers with long commutes. These findings are also consistent with the idea that efficiency wages tend to be paid to blue collar workers, who tend to be in occupations that face high levels of supervision presumably because the cost of shirking is high in those occupations.

Some previous research has examined the empirical relationship between commutes and labor market outcomes. In the spatial mismatch literature, a substantial number of studies use average commute time as a proxy for employment access and sometimes find a positive relationship between commutes and employment interpreting these findings as evidence of imperfections in urban labor markets.³ Further, Zax and Kain (1991) examined the quit rates of white and black employees following the relocation of their current employer from downtown Detroit to the suburbs. While they find that the change in commute times affected black quit rates, these changes had no effect on white employees who presumably faced less location constraints in the housing market. On the other hand, empirical studies of wages and commutes usually focus on the prediction in competitive labor markets that firms located further from workers, usually near the center of the central business district, must pay a wage premium to compensate workers for that longer commute. White (1999) provides a detailed survey of empirical studies of the urban wage gradient. Some examples include Madden (1985) who finds a positive relationship between wage change and change in commute for workers who changed jobs, Timothy and Wheaton (2001) who find that average commutes in an employment location can explain across location wage differences, and Manning (2003) who shows that an extra hour of commuting each day is associated, on

³However, studies employing more precise measures of employment access tend to find a more robust relationship (Ihlanfeldt and Sjoquist, 1998).

average, with an increase in wages of between 3 and 28 percent depending upon the sample and the specification. On the other hand, Zax (1991) finds a positive relationship between commutes and wages using the relocation of a single Detroit firm, which is the first evidence to suggest that wages may vary across workers at the same firm based on those worker's commutes, a phenomenon that is not consistent with competitive labor markets. Our theoretical model provides one possible explanation for why employment and wages might vary with commutes even after controlling for employment location, and the empirical portion of our study develops an instrumental variables strategy that allows us to test whether commutes or expected commutes as a result of a worker's residential location affects employment and wages.

Existing empirical studies of efficiency wages tend to focus on across industry wage differentials. Specifically, a wide variety of studies, Krueger and Summers (1987, 1988), Dickens and Katz (1987), Murphy and Topel (1987, 1990), and Gibbons and Katz, test for the existence of unexplained wage differentials and whether those differentials can be explained by unobserved differences in worker ability or productivity. Neal (1993) attempts a more direct test of the theory by examining whether across industry differences in supervision can explain across industry wage differentials. More recently, however, several studies focus on within industry or firm differentials with Chen and Edin (2002) examining wage differentials between piece rate and hourly pay and Lazear (2000) and Paarsh and Shearer (2000) examining the link between productivity and wages. This study complements these earlier studies by providing a test for efficiency wages that exploits an entirely different source of empirical variation; namely, the empirical relationship between employment and wages and a length of a worker's expected commute.

The remainder of the paper is organized as follows. The next section presents the theory. Sections 3 and 4 are devoted to the empirical strategy, the data and the results. In section 5, we perform some robustness checks. Finally, section 6 concludes.

2 Theory

2.1 The model

There is a continuum of workers (employed or unemployed) uniformly distributed along a monocentric, linear and closed city who endogenously decide their effort level at work e and the optimal residential location between the business district and the city fringe. They all consume the same amount of land (normalized to 1 for simplicity) and the density of

residential land parcels is taken to be unity so that there are exactly x units of housing within a distance x of the business district.

All firms are assumed to be exogenously located in the Business District (BD hereafter). The BD is a unique employment center located at one end of the linear city. In a centralized city, it corresponds to the central business district, whereas in a completely decentralized city, it represents suburban employment. As will be clear below, what is crucial here is not the location of the BD but the distance between workers' residential location and their workplace (i.e. the BD). All land is owned by absentee landlords.⁴ Each worker (employed or unemployed) who consumes one unit of land is assumed to be infinitely lived and risk neutral. Workers endogenously decide their optimal place of residence between the BD (i.e. 0) and the city fringe (x_f). The total population is normalized to 1 so that the unemployment rate is equal to the unemployment level and is given by u , Similarly, the employment rate is equal to the employment level and is given by $1 - u$.

At any moment, workers can either be employed or unemployed. If employed he/she obtains a wage w whereas if unemployed he/she gets an unemployment benefit b . We assume that changes in the employment status (employment versus unemployment) are governed by a continuous-time Markov process. Job contacts (that is the transition rate from unemployment to employment) randomly occur at an endogenous rate θ while the exogenous job separation rate is δ . In this context, the expected duration of employment is given by $1/\delta$ whereas the expected duration of unemployment amounts to $1/\theta$. It then follows that a worker spends a fraction $\theta/(\theta + \delta)$ of his/her lifetime employed and a fraction $\delta/(\theta + \delta)$ of his/her lifetime unemployed. In steady state, flows into and out of unemployment are equal. Therefore, we have:

$$u = \frac{\delta}{\theta + \delta} \tag{1}$$

Observe from (1) that the steady state unemployment and employment rates correspond to the respective fractions of time a worker remains unemployed and employed over his/her infinite lifetime. Equation (1) can also be interpreted as the probability a worker will be unemployed in steady state.

Let us now determine the instantaneous utilities of an employed and an unemployed

⁴All these assumptions are very standard in urban economics, see Fujita (1989). The key assumption that might give readers pause is the concentration of employment in one location. Many urban models have generalized the classic monocentric model to allow for decentralized or multi-centric employment, see Ross and Yinger (1995) for example. The main behavioral results arising from the monocentric urban model invariably hold in these more general model with individuals who face the same commuting costs making similar decisions regardless of their employment location.

worker. For the employed, the utility function is separable and is given by:⁵

$$z_1 + V(l, e)$$

where z_1 is the quantity of a (non-spatial) composite good (taken as the numeraire) consumed by the employed and $V(\cdot)$ is assumed to be increasing in leisure l and decreasing in effort e , and concave in both arguments. This choice of the utility function aims at capturing the fact that *effort and leisure are not independent activities*. Indeed, if one interprets $-e$ as the leisure activity on the job (shirking), then the benefits arising from additional *leisure activity on the job* is obviously related to the extent of *leisure activity at home* and vice versa.⁶ We assume that effort and leisure are substitutes,⁷ that is

$$\frac{\partial^2 V(l, e)}{\partial l \partial e} > 0 \quad (2)$$

Indeed, low leisure at home may imply that the worker has less time for rest and relaxation and more pressed for time at home (less time for relaxation or errands), and as a result the benefit of taking leisure or conducting home production (relaxation or errands) while at work rises. This story is consistent with increasing disutility from e as leisure falls or $\partial^2 V(l, e) / \partial l \partial e > 0$, and the level of e will fall as commutes increase.

We are now able to write the budget constraint of an employed worker. Each worker purchases the good z produced and incurs τx in monetary commuting costs when he/she lives at distance x from the BD. Letting $R(x)$ denote rent per unit of land, the budget constraint of an employed worker at distance x can be written as follows:

$$wT = z_1 + R(x) + \tau x \quad (3)$$

where w is the per-hour wage and T denotes the amount of working hours. T is assumed to be the same and constant across workers, an assumption that agrees with most jobs in the vast majority of developed countries.

⁵Subscripts ‘1’ and ‘0’ respectively refer to the employed and the unemployed groups.

⁶While this model does not incorporate a full home production model as in Becker (1965), leisure time at home might reasonably be viewed as a composite good that encompasses a variety of home activities. The traditional home production model implies that changes in the wage rate will cause substitution between home and market production (Baxter and Jerman, 1999). It also seems reasonable that changes in the time available at home, possibly due to a long commute, is likely to influence the distribution of personal activities between home and work.

⁷The case where effort and leisure are complements is not treated here because it is not consistent with what we find in the empirical analysis. However, in our working paper (Ross and Zenou, 2007), we have a complete analysis of this case.

Each worker provides a fixed amount of labor time T so that the time available for leisure l depends solely on commuting time. Thus, denoting by tx the commuting time from distance x (where $t > 0$ is the time commuting cost per unit of distance), the time constraint of an employed worker at distance x can be written as:

$$1 - T = l + tx \quad (4)$$

in which the total amount of time is normalized to 1 without loss of generality.

By plugging (3) and (4) into the utility function, we obtain the following indirect utility for the employed:

$$I_1(x, e) = z_1 + V(l, e) = wT - R(x) - \tau x + V(1 - T - tx, e) \quad (5)$$

Let us now focus on the unemployed. Their budget constraint is given by:

$$b = z_0 + R(x) + \tau x \quad (6)$$

where b is the unemployment benefit. In this formulation, we assume for simplicity that employed and unemployed workers have the same monetary commuting costs. The former commute every day to work whereas the latter commute every day for interviews.

To keep the analysis manageable and to be consistent with the utility of the employed, we assume that the unemployed's utility function is given by: $z_0 + V_0$,⁸ and, without loss of generality, we normalize b to zero. In this formulation, V_0 is a constant utility benefit arising to all who are unemployed. Basically, V_0 is introduced to recognize the inherent disutility to being at work and commuting to work since it assures that when people have exactly the same z , the one working can receive less utility.

By using (6), we obtain the following indirect utility function for the unemployed:

$$I_0(x) = z_0 + V_0 = -R(x) - \tau x + V_0$$

We are now able to calculate the expected utility of each worker. To do that, we assume perfect capital markets with a zero interest rate.⁹ Borrowing and lending are not permitted.

⁸This formulation assumes that there is no search behavior from the unemployed. They just obtain a job randomly. This is consistent with the standard assumptions of exogenous reemployment probability in the efficiency wage model. Observe that all our basic results go through if we allow for search with time and commuting costs for the unemployed. The analysis just gets messier.

⁹When there is a zero interest rate, workers have no intrinsic preference for the present so that they only care about the fraction of time they spend employed and unemployed. Therefore, the expected utilities are not state dependent.

We also assume that there are very high mobility costs.¹⁰ This implies that *a worker's residential location remains fixed as he/she enters and leaves unemployment*. This is much more realistic than assuming that changes in employment status involve changes in residential location. In fact, for workers to stay in the same location and thus pay the same bid rent over their lifetime, it has to be that they adjust their composite consumption. It is easy to verify that

$$z_1 - z_0 = wT > 0$$

which means that workers consume less composite good when unemployed. This difference increases with wages since better paid workers consume more composite good only when employed.¹¹

Since a worker spends a fraction $1 - u = \theta/(\theta + \delta)$ of his/her lifetime employed and a fraction $u = \delta/(\theta + \delta)$ unemployed, at any moment of time, the utility of a worker is thus equal to that worker's average utility over the job cycle and is given by

$$\begin{aligned} I &= (1 - u)I_1(x, e) + uI_0(x) \\ &= (1 - u) [wT + V(1 - T - tx, e)] - R(x) - \tau x + uV_0 \end{aligned} \quad (7)$$

Because workers can provide different effort levels depending on their location, the information available to firms about workers' residence will matter in the process of wage formation. As a result, we consider two different information structures. In the first one, firms cannot observe workers' residence while they can perfectly observe the residential location of all workers in the second case.

2.2 The equilibrium when firms do not observe workers' location

Each individual supplies one unit of labor. There are only two possible effort levels: either the worker shirks, exerting effort $e^S = \underline{e} > 0$, and contributing \underline{e} units to production, or he/she does not shirk, providing effort $e^{NS} = \bar{e} > \underline{e}$, and contributes \bar{e} units to production. The implication of this efficiency model, which allows for substitution between leisure and shirking behavior, differs from the standard efficiency wage model (Shapiro and Stiglitz, 1984) in that some shirking is possible and can persist in equilibrium.

¹⁰A similar assumption can be found in Brueckner and Zenou (2003) and Zenou (2006).

¹¹Many intertemporal models recognize that households might engage in precautionary savings in order to protect against negative shocks, such as unemployment, and in those types of models consumption may not fall or at least will fall less during spells of unemployment. In this model, however, consumers have no incentive to smooth consumption because consumption of the composite commodity, z , enters utility in a linear fashion and interest and discount rates are both zero.

Using (7), and given that all workers obtain the same wage, this implies that the expected indirect utilities of non-shirker and shirker workers are respectively equal to:¹²

$$I_{no}^{NS}(x, \bar{e}) = (1 - u^{NS}) [w_{no}T + V(1 - T - tx, \bar{e})] - R_{no}(x) - \tau x + u^{NS} V_0$$

$$I_{no}^S(x, \underline{e}) = (1 - u^S) [w_{no}T + V(1 - T - tx, \underline{e})] - R_{no}(x) - \tau x + u^S V_0$$

which are simply weighted averages of the utility levels when employed and unemployed and where the share of time spent unemployed is used to form the weights.

Since shirking is not perfectly detected by firms, we assume that there is a monitoring technology m (probability of detecting shirking).¹³ Using (1), this implies that

$$u^{NS} = \frac{\delta}{\theta + \delta} \quad (8)$$

$$u^S = \frac{\delta + m}{\theta + m + \delta} \quad (9)$$

with $u^S > u^{NS}$, $\forall \delta, \theta, m > 0$. All workers (shirking or not shirking) must in equilibrium obtain the same utility level \bar{I}_{no} , which is location independent. Since workers stay in the same location all their life, bid rents are given by:¹⁴

$$\Psi^{NS}(x, \bar{I}_{no}) = (1 - u^{NS}) [w_{no}T + V(1 - T - tx, \bar{e})] - \tau x + u^{NS} V_0 - \bar{I}_{no} \quad (10)$$

$$\Psi^S(x, \bar{I}_{no}) = (1 - u^S) [w_{no}T + V(1 - T - tx, \underline{e})] - \tau x + u^S V_0 - \bar{I}_{no} \quad (11)$$

Inspection of these two equations shows that, as usual, the bid-rent functions are decreasing in x , with $\partial \Psi^{NS}(x, \bar{I}_{no}) / \partial x < 0$ and $\partial \Psi^S(x, \bar{I}_{no}) / \partial x < 0$. In the present model, this reflects the combined influence of the time cost of commuting and the monetary cost. Let us denote by \tilde{x} the border between non-shirkers and shirkers. We have the following result.¹⁵

¹²The subscript *no* refers to the “non-observable” case.

¹³In the present model, as in most efficiency wage models, firms can choose the wage w , but have no ability to change monitoring effort m . It is just assumed that monitoring and wages are substitute instruments. However, by developing a model where firms choose both w and m , Allgulin and Ellingsen (2002) have shown that, if workers only have a choice between two effort levels (working and shirking), the accuracy of monitoring and the level of pay are still negatively related.

¹⁴The use of bid-rent curves are standard in models with land markets and commuting, and in these types of models rents are assumed to adjust so that in equilibrium consumers are indifferent between different locations in the city.

¹⁵All proofs of propositions can be found in the Appendix.

Proposition 1 *Assume (2). Then, workers who reside close to jobs will choose not to shirk whereas workers located farther away will shirk.*

Indeed, if effort and leisure are substitutes, workers residing close to jobs will provide more effort than those residing further away from jobs because they have lower commuting time and thus more leisure time at home. Let us now determine \tilde{x}_{no} , the boundary between shirking and non-shirking workers where a consumer is indifferent between high and low levels of effort at work. To obtain the value of \tilde{x}_{no} , we have to solve: $\Psi^{NS}(\tilde{x}_{no}, \bar{l}_{no}) = \Psi^S(\tilde{x}_{no}, \bar{l}_{no})$, which is equivalent to:

$$(1 - u^S)V(1 - T - t\tilde{x}_{no}, \underline{e}) - (1 - u^{NS})V(1 - T - t\tilde{x}_{no}, \bar{e}) = (u^S - u^{NS})(w_{no}T - V_0) \quad (12)$$

showing a clear trade off between shirking (higher utility $V(\cdot)$ when employed since less effort but more unemployment spells during the lifetime) and nonshirking. Adopting the following notations, $V(1 - T - t\tilde{x}_{no}, \bar{e}) \equiv \tilde{V}_{no}^{NS}$ and $V(1 - T - t\tilde{x}_{no}, \underline{e}) \equiv \tilde{V}_{no}^S$, we have:

$$\frac{\partial \tilde{x}_{no}}{\partial w_{no}} = \frac{T(u^S - u^{NS})}{t \left[(1 - u^{NS}) \frac{\partial \tilde{V}_{no}^{NS}}{\partial l_{no}} - (1 - u^S) \frac{\partial \tilde{V}_{no}^S}{\partial l_{no}} \right]} \quad (13)$$

Proposition 2 *Assume (2). Then, higher wages reduce the fraction of nonshirkers in the city, i.e. $\partial \tilde{x}_{no} / \partial w_{no} > 0$.*

This proposition states that wages affect \tilde{x}_{no} the border between shirkers and non-shirkers. Indeed, if (2) holds, i.e. effort and leisure are substitutes, then when wages are higher, less workers shirk (the fraction of shirkers $1 - \tilde{x}_{no}$ decreases) since there are more incentive not to shirk (the average wage difference $w_{no}T(u^S - u^{NS})$ between shirkers and non-shirkers increases).

Lemma 1 *Assume (2) and*

$$\frac{\partial^2 \tilde{V}_{no}^{NS}}{\partial l_{no}^2} < \frac{\partial^2 \tilde{V}_{no}^S}{\partial l_{no}^2} < 0 \quad (14)$$

Then, $\partial^2 \tilde{x}_{no} / \partial w_{no}^2 < 0$.

To obtain this result, the main condition is (14) because it guarantees an interior solution by separating workers over space. The intuition behind this assumption is quite reasonable and consistent with the intuition behind the assumption stated in equation (7). Consider a plot of the marginal utility of leisure $\partial V / \partial l$ against effort with effort on the horizontal axis,

which is positively sloped by equation (7). For low levels of the marginal utility of leisure (high levels of leisure), effort at work probably has little impact on the marginal utility of leisure because the worker is well rested and his/her home is well ordered. The resulting plot of $\partial V/\partial l$ is fairly flat. On the other hand, for high levels of $\partial V/\partial l$ (low leisure), effort at work is probably quite important, and the plot of $\partial V/\partial l$ is likely to be quite steep. Under these conditions, the effect of a decrease in l on $\partial V/\partial l$ is larger in magnitude at high levels of effort, which is consistent with equation (14).

Let us now determine the labor market equilibrium. We consider a closed city model in which \bar{l}_{no} is endogenous and the city fringe is equal to 1 (the size of the total population is 1 since land consumption is 1). There are M_{no} firms in the economy. The profit function of a typical firm can be written as:

$$\Pi = F(\alpha_{no} [L_{no}^{NS} \bar{e} + L_{no}^S \underline{e}]) - w_{no} \alpha_{no} L_{no}$$

where the total number of employed workers is $L_{no} = L_{no}^S + L_{no}^{NS}$, α_{no} is the fraction of workers hired by each firm, and $F(\cdot)$ is the production function, with $F'(\cdot) > 0$ and $F''(\cdot) < 0$. The total number of non-shirkers in the economy is given by

$$L_{no}^{NS} = \tilde{x}_{no}(1 - u^{NS}) \quad (15)$$

the total number of shirkers is

$$L_{no}^S = (1 - \tilde{x}_{no})(1 - u^S) \quad (16)$$

We impose here that there is no discrimination in wages, which means that all workers, whatever their location, obtain the same wage. We also impose that firms employ the same fraction α_{no} of workers (shirkers and nonshirkers).¹⁶ Thus, even if firms know that all workers residing beyond \tilde{x}_{no} will shirk, they have to pay them the same wage as the ones who live between 0 and \tilde{x}_{no} (nonshirkers). This, in particular, implies that *employment is lower when commuting times are longer*.

Let us now solve the firm's program. By taking \bar{e} , \underline{e} , u^S and u^{NS} as given, each firm chooses w_{no} and α_{no} that maximize its profit. When choosing w_{no} firms will face the following trade off. Because it affects \tilde{x}_{no} , higher wages implies that the fraction of shirkers hired will be lower (Proposition 2) and thus total output increases but labor costs are also higher since the wage given to workers is the same. When choosing α_{no} firms face the following trade

¹⁶This might seem unreasonable, but actually if the location of the workers at a given firm are distributed randomly then over time the firms share of shirking workers at any wage should mirror the economies share, and individual firms can directly effect their fraction of shirkers with their wage.

off. Higher α_{no} means that more workers are hired, which implies that firms have a higher output but also higher labor costs. We have the following proposition.

Proposition 3 *Assume (2) and (14). Then, firms always want to allow some shirking in equilibrium and all shirkers live at the periphery of the city.*

We show that it is optimal for each firm to set a wage given by (30) in the Appendix. This wage is set by taking into account the fact that it affects \tilde{x}_{no} , the fraction of non-shirkers in the each firm, via (12). Of course, one has to verify that the wage, given by (30) in the Appendix, indeed maximizes firms' profit and corresponds to a strictly interior \tilde{x}_{no} , i.e. $\tilde{x}_{no} \in]0, 1[$. We assume here a strictly interior solution for \tilde{x}_{no} .

In equilibrium, it has to be that labor supply equals labor demand for nonshirkers and shirkers respectively. Since the total population of workers is equal to 1, these conditions can be written as:

$$\begin{aligned}\alpha_{no}M_{no}L_{no}^{NS} &= \tilde{x}_{no}(1 - u^{NS}) \\ \alpha_{no}M_{no}L_{no}^S &= (1 - \tilde{x}_{no})(1 - u^S)\end{aligned}$$

Using (15) and (16), this implies that

$$M_{no} = \frac{1}{\alpha_{no}} \quad (17)$$

We are now able to define the equilibrium of this economy. Assume that (2) and (14) hold. Then, nonshirkers are close to jobs whereas the shirkers are far away. In the Appendix, we give a formal definition of this equilibrium and it can then be shown that it exists and is unique.

2.3 The equilibrium when firms do observe workers' location

Assume now that firms do observe workers' location. The utility of each worker is still given by (7). Now, as we will show below, there will be no shirking in equilibrium. This implies that the unemployment rate of the economy is given by (8). The bid rent of a (non-shirker) worker is equal to:¹⁷

$$\Psi(x, \bar{I}_o) = (1 - u^{NS}) [w_o(x)T + V(1 - T - tx, \bar{e})] - \tau x + u^{NS} V_0 - \bar{I}_o \quad (18)$$

Inspection of this equation shows that

$$\frac{\partial \Psi(x, \bar{I}_o)}{\partial x} = (1 - u^{NS}) \left[w'_o(x)T - \frac{\partial V(1 - T - tx, \bar{e})}{\partial l_o} \right] - \tau \quad (19)$$

¹⁷The subscript o refers to the "observable" case.

which can be positive or negative depending on the sign of $w'_o(x)$ (it will be determined below). Since all workers provide the same effort level and are identical in all respects, they just locate anywhere in the city and enjoy the same utility level \bar{l}_o , the land rent adjusting for commuting cost differences between different locations.

To close the urban equilibrium, we have to check that $\Psi(1, \bar{l}_o) = 0$, which is equivalent to:

$$\bar{l}_o = (1 - u^{NS}) [w_o(1)T + V(1 - T - t, \bar{e})] - \tau + u^{NS} V_0 \quad (20)$$

At each location x in the city ($0 \leq x < 1$), each firm has to set a NSC (that equates shirking and non-shirking utilities) to prevent shirking. We obtain¹⁸

$$w_o(x) = \frac{(1 - u^S)V(1 - T - tx, \underline{e}) - (1 - u^{NS})V(1 - T - tx, \bar{e})}{T(u^S - u^{NS})} + \frac{V_0}{T} \quad (21)$$

This is the standard Shapiro-Stiglitz style non-shirking condition evaluated in equilibrium for every residential location x . It should be clear that, when firms can observe workers' location, it is optimal for them to wage discriminate in terms of location and not to allow shirking in equilibrium. In the previous model, this was not possible since each firm had to give to all its workers the same wage and thus it was somehow optimal to let some workers shirk.

Proposition 4 *Assume (2). Then, $w'_o(x) > 0$.*

Thus *wages are higher when commuting times are longer*. Indeed, if leisure and effort are substitutes (i.e. (2) holds), then wages have to compensate workers who live *further away* since they commute more and thus have less time for leisure at home. As a result, wages increase with distance to jobs.

Using (19), Proposition 4 implies that when (2) holds, $w'_o(x) > 0$ and thus the sign of $\partial\Psi(x, \bar{l}_o)/\partial x$ is ambiguous. This is because there are two opposite effects. On the one hand, workers residing far away have higher wages. On the other, they have higher monetary commuting costs and also higher time costs and thus lower leisure. The compensation of the land rent is therefore not straightforward. Because we would like land rents to decrease from the center to the periphery, we assume that

$$(1 - u^{NS})w'_o(x)T < (1 - u^{NS})\frac{\partial V}{\partial l_o}t + \tau$$

¹⁸As in the previous model, we have here the assumption that firms receive an equal share of workers from each location.

i.e., the wage is lower than the commuting cost effect so that land rents compensate workers who reside further away. Using (8) and (9), this can be written as:

$$0 < \frac{\partial V(1 - T - tx, \bar{e})}{\partial l_o} - \frac{\partial V(1 - T - tx, \underline{e})}{\partial l_o} < \frac{\tau m}{t\theta} \quad (22)$$

which guarantees that bid rents are always decreasing. This condition encompasses (2).

Let us now define the labor demand α_o of each firm. Since all firms are identical, let us focus on a symmetric (steady-state) equilibrium in which each firm employs the same number of workers α . The profit of each firm is given by:

$$\Pi_o = F(\alpha_o L_o \bar{e}) - \alpha_o L_o \int_0^1 w_o(x) dx \quad (23)$$

and thus firms solve the following program: $\max_{\alpha} \Pi_o$. First order condition yields

$$\bar{e} F'(\alpha_o L_o \bar{e}) = \int_0^1 w_o(x) dx \quad (24)$$

where, using (21), we have:

$$\int_0^1 w_o(x) dx = \frac{1}{T(u^S - u^{NS})} \left[(1 - u^S) \int_0^1 V(1 - T - tx, \underline{e}) dx - (1 - u^{NS}) \int_0^1 V(1 - T - tx, \bar{e}) dx \right] + \frac{V_0}{T}$$

Equation (24) means that each firm adjusts employment until the marginal product of an additional worker equals the average efficiency wage. Finally, the equilibrium condition (labor demand equals labor supply) is given by:

$$L_o = 1 - u^{NS} = \frac{\theta}{\delta + \theta} \quad (25)$$

In the Appendix, we give a definition of this equilibrium. Again, it can then be shown that it exists and is unique.

2.4 The intermediate case: Firms partially observe workers' location

The theory has discussed so far the two extreme cases where firms can or cannot observe where workers live. An intermediate case may arise if commute time is only partially observed. For example, firms may observe residential location, which provides a noisy signal of

a worker's commute, but it is too costly for firms to determine the true commutes of workers. Further, residing a long distance from one's workplace leads to both a longer commute and a greater variance in the error associated with a firm's estimate of a worker's commute because the long distance creates more opportunities for idiosyncratic factors that affect point to point commutes and because distance may limit a firm's knowledge of such idiosyncratic factors. Firms will pay efficiency wages to reduce shirking, but they will not know the exact premium required to eliminate shirking by the worker or workers residing at a given location. Presumably, firms will increase wages until the benefit of reducing the likelihood of shirking for a worker or workers equals the cost of the wage increase. Some workers are likely to shirk in equilibrium because their commute is substantially longer than the firm's estimate of their commute, e.g. they have a commute time draw that occurs with very low probability and so firms would not rationally raise wages to eliminate shirking under such a low probability event. As the noise in commutes increases at longer distances, the required increase in wages to eliminate a certain amount of shirking at a location increases, and firms rationally do not eliminate as much shirking (pushing wages up above the level paid to workers who live closer $dw/dx > 0$, but leaving shirking levels higher for larger values of x). Therefore, in the intermediate case, both the employment and wage conditions implied by propositions 3 and 4 should hold in equilibrium.¹⁹

3 Empirical strategy

The theoretical models above suggest an empirical relationship between commuting time and either job separation and/or wages. While the model only provides an unambiguous prediction concerning the sign of these relationships under a specific set of assumptions and other models might generate similar relationships, the model above does suggest that if efficiency wages are going to play an important role in the spatial distribution of unemployment or wages this role should be directly related to the commutes faced by workers (since, assuming (2), commutes and thus leisure have a positive impact on effort and thus shirking). Further,

¹⁹A second intermediate scenario, which allows commute time to be observed albeit at some cost, might arise because some firms wage discriminate based on workers' commutes eliminating all shirking and others simply pay a uniform efficiency wage to all workers possibly because the cost of monitoring commute time varies across firms. This scenario should also lead to a situation where both unemployment and wages increase systematically with distance, but the exact equilibrium configuration is complicated by the fact that firms that pay commute specific efficiency wages face an adverse selection problem because workers with a long commute and therefore a high likelihood of shirking are attracted to the firms that will pay them the high wages that prevent shirking.

we would expect these relationships to be strongest for individuals who are likely to be paid efficiency wages as an alternative to costly supervision. The tests offered in this paper can be viewed as necessary conditions, as opposed to sufficient, for efficiency wages to be important in the spatial distribution of labor market outcomes. Moreover, a finding that one of the variables, unemployment or earnings, are related to commutes, but not the other would provide evidence favoring the model where firms cannot or can discriminate over space when setting wages.

The direct inclusion of commute time in an employment or wage model results in a number of important biases. First, households choose their residential location in part based on demand for housing and neighborhood attributes, which is heavily influenced by the permanent incomes of the household members. Accordingly, unobservables that affect location choice are likely to be correlated with labor market outcomes. In addition, commute time is usually measured with significant error and is usually unobserved for unemployed individuals leading to measure error and selection bias, respectively. A standard solution for these biases is to estimate the effect of commute time using an instrumental variables specification, which requires that the effect of commuting time be identified by exclusion of some factor that can explain commuting time differences from the labor market equation

Cutler and Glaeser (1997) argue that an appropriate solution to the bias caused by sorting over residential location is to aggregate up to a level at which sorting is less likely to be endogenous, and they recommend identifying the effect of location on outcomes using across metropolitan variation. This approach has been used in a variety of applications including Evans, Oates, and Schwab (1992) on neighborhood poverty, Cutler and Glaeser (1997) on residential segregation, Ross (1998) and Weinberg (2000) on the spatial mismatch hypothesis, Hoxby (2000) on school competition, and Card and Rothstein (2005) on the black-white test score gap. In all of these papers, a cross-sectional analysis is conducted in which across metropolitan variation in a key variable, such as rivers, number of jurisdictions, or job decentralization, is used to identify the impact of location after controlling for other important metropolitan level variables.²⁰

We extend this approach by considering a broader source of variation for identification. Specifically, we recognize that a metropolitan area presents households with a complex set of location options and that the expected location attributes arising from this choice set is likely to vary systematically across individuals. For example, the limitations in choice created by a racially segregated metropolitan area likely vary between college and high school educated

²⁰For more general discussions of identification in neighborhood effect models, see Moffitt (2000), Bayer, Ross and Topa (2005), Durlauf (2004), and Bayer and Ross (2006).

African-Americans. We estimate a model of an individual's commute time (Z_{ils})²¹ and other location attributes based on a specific set of individual attributes (W_{ils}) and the individual's metropolitan area of residence. Specifically,

$$Z_{ils} = \beta_s W_{ils} + \varepsilon_{ils} \quad (26)$$

where β_s describes a relationship between the individual attributes and commute time that is unique to metropolitan area s . Labor market outcomes (Y_{ils}) are modeled as function of predicted commute time from equation (26), a broader set of individual attributes, X_{ils} and W_{ils} , as well as metropolitan area fixed effects, or

$$Y_{ils} = \gamma X_{ils} + \zeta W_{ils} + \delta E[Z_{ils}|W_{ils}, \hat{\beta}_s] + \alpha_s + \mu_{ils} \quad (27)$$

In this specification, no individual variables are excluded from the second stage. Rather, identification is accomplished by the restriction that the relationship between W_{ils} and Y_{ils} does not vary systematically across metropolitan areas, which is similar to the identification strategy in housing price hedonics where the price function varies across markets and preferences are restricted to be the same across markets (Epple, 1987). Specifically, across individual differences in W_{ils} do not provide any information to identify δ . Rather, across metropolitan area differences in the estimated values of β_s create differences in expected commutes for observationally equivalent individuals who reside in different metropolitan areas. A second and related feature of this model is that some variables are intentionally omitted from the first stage model for Z_{ils} . The impact of some individual attributes on Y_{ils} may vary across metropolitan areas or may be endogenous to across metropolitan differences in outcomes. A straightforward solution to this problem is simply to allow the coefficients on those variables to vary across metropolitan areas in the second stage equation, as well. This approach, however, may lead to very complex models and in the interest of parsimony it may be preferable to simply omit those variables from the first stage so that identification cannot arise from that specific set of variables.

Finally, unlike the across metropolitan papers cited earlier, the identification strategy in this paper generates substantially more information allowing for the inclusion of multiple location variables and metropolitan area fixed effects. These additional controls mitigate concerns that the estimated effect of commutes on labor market outcomes is biased by the omission of important location or metropolitan area controls. In contrast to earlier papers, which examine the impact of one variable on all individuals in the metropolitan area, this

²¹The subscripts i , l and s refer respectively to the individual, his/her location and the metropolitan area where he/she lives.

model allows the expected value of Z_{ils} to vary across households in the same metropolitan area generating within metropolitan variation in the variable of interest. This additional source of variation allows for identification of coefficients on multiple location variables, as well as for the inclusion of metropolitan area fixed effects.

The specific identification assumption used in this paper is that the spatial distribution of the housing stock has an impact of commutes and other location variables that vary across metropolitan areas, but that the housing stock variables only enter an individual's labor market equation for the sample as a whole. Specifically, commute time, as well as individual exposure to good employment access, high income households and other location attributes, is regressed on a set of housing variables representing the individual's housing consumption. An individual's location variables are predicted based on the type of housing unit in which they reside, and the spatial distribution of housing units across the metropolitan area creates variation in exposure to these location variables. While the housing stock variables are not expected to have a causal impact on labor market outcomes, they are included in the second stage model along with standard demographic controls for labor market models in order to assure that the coefficients on the location variables are only identified based on across metropolitan area variation. As discussed earlier, only systematic across metropolitan differences in the relationship between housing stock and commutes are used to identify the impact of commutes on employment outcomes. Across household differences in the housing stock selected are captured by the direct inclusion of housing controls into the labor market model.

4 Data and results

The Public Use Microdata Sample (PUMS) of the 2000 Decennial Census is used to examine this question empirically because this sample records individuals' labor market outcomes, commutes, detailed individual attributes including industry and occupation, a reasonable set of housing attributes, and identifies each individual's residential and employment location to be within a specific spatial zone consisting of 100,000 individuals or more, Public Use Microdata Areas (PUMA's). A sample of 30 metropolitan areas are selected that have at least 5 residential PUMA's and 3 workplace PUMA's or at least 4 residential and 4 workplace PUMA's (at least 15 residential to workplace PUMA combinations). For our employment analysis, we draw a sample of all prime-age males (ages 25 to 55) who were either employed the week prior to the census survey or describe themselves as actively seeking work during that period (labor force participants), who do not reside in groups quarters, leading to a

sample of 1,127,607 workers. The wage rate is calculated as simply annual 1999 earnings divided by the product of typical hours worked per week last year times the number of weeks worked last year. The model to explain the logarithm of wages is estimated for all workers who worked at least 40 weeks last year and at least 30 hours in a typical week and whose estimated wage exceeded the federal minimum wage of \$5.15 per hour leading to a sample of 903,255 full-time workers.²²

In order to test whether the relationship between commutes and labor market outcomes are related to the operation of efficiency wages, we divide workers into 12 major occupational categories and classify those categories as lightly or heavily supervised.²³ This supervision classification is based on an analysis by Levenson and Zoghi (2006). They order occupations on a white collar-blue collar continuum with executives and managers at one end and unskilled laborers at the other, use data from the 1999 National Compensation Survey to plot average supervision levels within occupations, and fit these averages to a quartic trend line.²⁴ The estimated trend line identifies a clear break in the pattern of supervision with all major occupation categories on the white collar side of a threshold having an predicted supervision level between 0.62 and 0.66 and all categories on the blue collar side having a predicted level between 0.34 and 0.45 on a scale between zero and one where one implies independence from supervision. The sample is then split into sets of workers in high and low supervision occupations with approximately 62.6 and 60.4 percent of workers in the employment and log wage samples, respectively, in high supervision occupations.²⁵

²²Alternative models were estimated using a smaller sample of prime age male workers who report exactly 40 hours of work per week, and the empirical results are qualitatively similar. This sample restriction might be more consistent with the assumption of fixed work hours in the theoretical model, but was not used for the analysis presented in the paper because that sample is much more highly selected than the sample of worked 30 hours or more, prime age male workers.

²³The twelve major categories are manager, engineer and scientist, education and other professionals, Health professionals, Health support and protective services, food and personal services, sales, office, construction, maintenance, production, and transportation.

²⁴The 1999 National Compensation Survey (NCS) is a restricted use data set and is the first nationally representative survey to collect information on worker supervision since a 1977 supplement to the Panel Survey of Income Dynamics used in Neal (1993). The authors do not have access to the NCS data, and the U.S. Department of Labor has not published any descriptive statistics or tabular information on supervision from the survey.

²⁵The reader should note that the resulting estimates after splitting the sample by whether the individual's occupation is highly supervised or not are quite robust over a variety of model specification not shown in the paper; including models based on the dependent variables earnings and whether employed rather than the wage and employment conditional on labor force participation that are used in this paper and more relevant to our theoretical model. On the other hand, the estimated coefficient on commute time

An important question to address is whether we should expect to find efficiency wages operating in lightly or heavily supervised occupations. Occupations might have low levels of supervision because they pay efficiency wages and therefore workers do not require close supervision. This fairly standard logic is consistent with the idea that it is hard to observe the performance of white collar workers and so those workers must be paid efficiency wages in order to assure that they do not shirk, while blue collar workers can be prevented from shirking by intensive monitoring. On the other hand, occupations in which supervision is very important and shirking a large problem may have high levels of supervision on average, but there also may be many firms that hire workers in those occupations and rely on paying wages above marginal productivity in order to encourage effort on the part of those workers. Further, for high skill occupations with low supervision, firms may rely on alternative methods for motivating workers and assuring high performance levels. For example, certain high skill occupations may provide long term, tournament style rewards in terms of a steep wage trajectory based on successful long term performance, which can be verified without active monitoring of short or medium term performance. If so, occupations with this type of wage structure would be unlikely to use extensive employee monitoring or pay efficiency wages. It is notable that several recent efficiency wage studies, which find strong evidence of efficiency wages, have explicitly focused on workers involved in industrial production where supervision is high and the workers studied are clearly blue collar workers.²⁶

The employment and earnings model specifications will include commute time along with a standard set of individual level control variables.²⁷ The controls include age and square of the worker's age, educational attainment dummy variables going from high school graduate to recipients of master's degree or higher, race and ethnicity controls, marital status and presence of children in the household, and whether born in U.S. and if not a control for time residing in the U.S. The model also includes industry fixed effects and fixed effects for each occupation-metropolitan area combination. The models are estimated separately for high changes substantially and is not robust across different specifications when the sample is pooled across all occupations, presumably because the model that determines employment and wages varies across highly and lightly supervised occupations.

²⁶See for example Chen and Edin's (2002) study of workers in the metalworking industries in Sweden, Paarsh and Shearer's (2000) study of workers at a tree planting firm in British Columbia, and Lazear's (2000) study of workers at the Safelite Glass Corporation in the United States.

²⁷The model controls and instruments for the individual's actual commute time. Alternative models were estimated using average commute time between the individual's PUMA of residence and PUMA of work. The results with that variable are very similar to those presented in the paper. This suggests that individual idiosyncrasies in mode choice do not drive our results, but systematic differences in mode choice across residential locations may play a role in the relationship between commute times and labor market outcomes.

and low supervision occupations using instrumental variable analysis where an instrument for commute time is developed by estimating metropolitan specific models of individual commute times as a function of the individual’s housing attributes. The means and standard errors of all variables are shown separately for workers in highly and lightly supervised occupations in Table 1. Workers in occupations that are highly supervised have slightly lower rates of employment, lower earnings, substantially less education, and have greater representation of minority and immigrant groups. As designed, the occupational split in the sample accurately distinguishes between white collar workers who tend to operate more independently and blue/pink collar workers who tend to be subject to more direct supervision.

[Insert Table 1 here]

The employment model is estimated using a linear probability model where observations are deleted iteratively whenever predicted probabilities from the estimated model fall above one or below zero. Horace and Oaxaca (2003) show that this iterative approach provides consistent estimates for binary dependent variable problems.²⁸ The wage model is estimated using standard ordinary least squares where the dependent variable is the logarithm of earnings. In both cases, a two stage approach is used where the predicted value of commute time and potentially other location attributes are included directly in the employment and wage equations. The housing variables used to instrument for commute time and other location attributes include number of bedrooms, number of other rooms, age of structure, number of units in structure, lot size, owner-occupancy, non-owner-occupant without cash rent, condominium, mobile home, incomplete kitchen facilities, and incomplete plumbing.²⁹ All standard errors are clustered at the occupation-metropolitan area level in order to address the concern that standard errors might be biased due to heteroscedasticity across occupations within a specific metropolitan area or across metropolitan areas for the same occupation.

Table 2 presents the results for the employment estimations. The estimates of the standard demographic controls are as expected with higher levels of education leading to higher employment levels, family structure influencing employment in the standard manner, minority groups having lower levels of employment, and immigrants having higher levels of employment. The estimates differ between the two occupation groups with the impact of race, family structure, and immigrant having substantially less influence on employment for the low supervision/white collar occupations. In terms of commutes, workers in high

²⁸The estimation of employment models using a logit yields qualitatively similar results.

²⁹The housing variables number of bedrooms, number of other rooms, age of structure, number of units in structure, and lot size are all represented by a series of dummy variables representing specific numbers of rooms or ranges for age, number of units, and lot size.

supervision occupations/blue collar workers have a negative relationship between commute and employment with a one hour increase in commute reducing employment rates by 2.4 percentage points. An increase in commutes reduces the time endowment of these workers and appears to increase the likelihood of unemployment potentially because the likelihood of shirking has risen for these workers. This finding is consistent with our assumption (2) of a positive cross-partial derivative of utility with respect to effort and leisure, or intuitively substitutability between leisure time and shirking. The coefficient on commutes for low supervision/white collar workers is near zero. It also is notable that the inclusion of and instrumenting for commute time has no impact on the parameter estimates for the demographic controls.

[Insert Table 2 here]

Table 3 presents the results for the wage estimations. The estimates are similar to the estimates for the employment model with the estimates on the demographic variables following standard patterns for labor market models. As before, the estimated impact of family structure and being an immigrant are substantially smaller for occupations with low levels of supervision, and the estimates on demographic variables are unaffected by the inclusion of and instrumenting for commute time. The effect of commute for workers in highly supervised occupations is consistent with efficiency wage based discrimination using commute time where a one hour increase in commute leads to a 8.2 percent increase in wages. The commute time effect on wages also implies a positive cross-partial derivative between effort and leisure, which again corresponds to assumption (2). These findings are consistent with incomplete commute based wage discrimination so that the unemployment effect of commutes through increased shirking remains and is captured by the estimations presented in the first column of Table 2.

Observe that our results are in line with our efficiency wage model and not with a standard search approach. Indeed, theoretically, long distance search (presumably correlated with a long expected commute) may lower the effectiveness of search and raise the cost. These two effects combine to reduce the number/quality of offers received and lower the reservation wage yielding an ambiguous effect on the likelihood of employment effect. However, lower quality offers and lower reservation wages both should lead to lower wages on average for people with long commutes. This theoretical prediction is the opposite of the empirical findings in our model where we find that long commutes lead to higher wages presumably to reduce shirking.

[Insert Table 3 here]

On the other hand, the estimations indicate a negative impact of commutes on wages for workers in lightly supervised occupations, and while smaller in magnitude than the results for highly supervised workers the estimate is statistically significant at a 95 percent level of confidence. In principle, these results might be consistent with white collar workers having a negative cross-partial derivative of utility with respect to effort and leisure, and firms being capable of perfect wage discrimination for white collar workers so that commutes have no influence on employment for this group. However, we do not believe that such strong conclusions are warranted. First, the negative result on commutes for lightly supervised workers is not entirely robust, as will be seen in the next section.

Second, a more conservative interpretation of these results is as a test for differences in the effect of commutes on employment and wages between workers in lightly and highly supervised occupations. In that context, we see that an hour increase in commute lowers employment in highly supervised occupations relative to those in lightly supervised occupations by 2.2 percentage points, and raises wages in the highly supervised occupations by 13.7 percent relative to lightly supervised occupations. While this interpretation does not allow us to sign the cross-partial derivative, the results provide strong evidence that commutes affect employment and wages differently based on the level of supervision in the occupation, and the direction of the effects on employment and wages are consistent with each other given the theory developed earlier.

5 Robustness checks

5.1 Omitted location attributes

This first set of robustness tests involve the inclusion of additional controls for the impact of residential location on neighborhood outcomes. For example, average commute time has sometimes been used as a proxy for employment access in studies of the spatial mismatch hypotheses. While the existing literature has established that commute time is a very poor proxy for employment access,³⁰ any correlation between commute time and employment access may bias our analysis unless employment access is explicitly included in the model. Similarly, the demographic composition of a worker's residential location may both influence labor market outcomes and be correlated with average commute times. Since workers

³⁰In practice, the correlation between gravity based measures of employment access and mean commute time is small. Furthermore, mean commute time is often insignificant in those studies while gravity based measures are typically found to be significant in the same or similar samples. See Ihlanfeldt and Sjoquist (1998) for a survey of the spatial mismatch literature.

select their residential locations and their likely commutes simultaneously, the instrumental variables estimation strategy will be applied to commute time as well as all other location variables.

Tables 4 and 5 present the results of series of analyses that included employment access,³¹ average family income, and neighborhood demographics into the employment and wage models, respectively. The inclusion of employment access into the model has no impact on the estimated effect of commutes on employment for occupations with supervision, but the inclusion of average family income reduces the magnitude of the parameter estimate on commutes by half and the commute time coefficient is even smaller and statistically insignificant after including the neighborhood demographic controls. On the other hand, Table 5 shows that the positive effects of commutes on the earnings of workers in occupations with high levels of supervision are very robust to the inclusion of any and all additional location variables. The finding of no employment effect for commutes for occupations without supervision is robust (Table 4 second panel), and the small negative earnings effect of commutes for occupations without supervision disappears after controlling for employment access (Table 5 second panel) bringing the employment and earnings estimates in line with each other for this group of workers.

[Insert Tables 4 and 5 here]

One possible explanation for the decline in the magnitude of the commute time coefficient is that the instrumental variables do not contain sufficient variation to identify the coefficients for so many endogenous regressors. This problem is analogous to the weak instrument problem for the single variable instrumental variables model. The standard rules of thumb for determining whether instruments are weak are well developed for the single endogenous variable case, but are not well defined for more complex models. However, a widely accepted feature of IV estimations based on weak instruments is that the IV estimates are biased towards the OLS estimates (Bound, Jaeger, and Baker, 1995; Zivot, Startz, and Nelson,

³¹The measure of employment access is created by estimating a traditional gravity model based on commuting flows between residential and workplace PUMA's. Specifically, the flow of commuters between each residential and workplace PUMA is regressed upon the number of prime age adults residing in the PUMA, the number of employed adults working in the PUMA, and the mean travel time between the residential and workplace PUMA using a log-log specification. The employment access measure is based on weighted exponential average of zone employment totals, where the parameter estimates and the commute times between a tract and each zone are used to create the weighting scheme. The employment access variable is then standardized so that the estimated coefficient represents the influence of a one standard deviation improvement in employment access. See O'Regan and Quigley (1996) for details on gravity based and alternative employment access.

1998).

In this spirit, we re-estimate all four models in Table 4 for workers in highly supervised occupations and find that the OLS estimates on commute time are near zero for all models. This finding is consistent with the hypothesis that the commute time estimates in the top panel of Table 4 are trending towards zero as more endogenous regressors are added due to a weak instruments bias. As a further test, we estimate a model that includes all location controls, but only instruments for commute time. In this model, the commute time estimates are statistically significant and close to the IV estimates in the first two columns. For example, for the full specification shown in the last column of Table 4, we find that the OLS estimate for commute time is 0.001, but the IV estimate for commute time when the weak instrument problem is avoided by using the actual value for other location controls is -0.018, which is smaller than, but close to the estimates in the first two columns of Table 4. The reader should also note that the OLS estimate for commute time on wages for workers in highly supervised occupations is 0.024 or near zero differing substantially from the IV estimates in Table 5 and suggesting that the wage models do not suffer from a serious weak instruments problem. Further, replacing the predicted value of all other location controls with actual values and continuing to instrument for commute time yielded estimates of 0.130, which are very similar to the IV estimates in the top panel of the last column in Table 5.

5.2 Exogeneity of instruments

A second concern arises from the possible endogeneity of the instruments. The most important control for limiting the bias from endogeneity is the inclusion of the housing stock variables in the labor market models. These controls assure that the estimates on commute time and other location variables are only identified based on across metropolitan differences in the distribution of households across each metropolitan area's housing stock. However, the identifying across metropolitan variation still might be endogenous to employment outcomes. For example, the unobservable, average human capital differences between owner-occupants of traditional single-family housing may differ systematically across metropolitan areas due to differences in market conditions.³²

In order to address this concern, we identify a subset of the housing stock variables, which we believe are most likely to be correlated with unobserved determinants of labor market outcomes, and re-estimate the IV model excluding those variables. Intuitively, this robustness

³²See Deng, Ross and Wachter (2003) for evidence of a simultaneity between homeownership and residential location choice.

check is analogous to an over-identification test where if the excluded instrumental variables are correlated with unobservables in the second stage labor market equation its exclusion from the first and second stage in this analysis should change the estimated coefficient on the commute time variable. For our application, variables that relate to owner-occupancy (owner-occupancy, non-owner-occupant without cash rent, condominium, mobile home) and variables that suggest that the housing might be substandard (incomplete kitchen facilities and incomplete plumbing) are deleted from both the first and second stage models under the assumption that across metropolitan differences in the spatial distribution of owner-occupied housing and substandard housing is less clearly exogenous than the spatial distribution of physical space variables, such as number of rooms or lot size.

Tables 6 and 7 provide the results for these analyses. The results are very robust the elimination of the owner-occupancy and the substandard housing variables has no meaningful impact on the point estimates for any of the specifications considered in either the employment or wage models. The findings presented in this paper appear to be quite robust to the choice of instruments.

[Insert Tables 6 and 7 here]

5.3 Role of education

Finally, the low supervision occupation workers have much more education on average than the high supervision workers, with 37.7 percent of the low supervision sample having a four year college degree compared to only 13.0% of the high supervision sample. Given the important human capital differences between the low and high supervision occupations, we interact education with commute time in order to see whether the supervision level differences in the effect of commutes are simply attributable to differences in the education level of the workers in the two sample. If education differences are truly responsible for our findings, the interaction of commute time with education level should eliminate the supervision level differences in the effect of commute time on labor market outcomes, and this extended model should simply find education differences in the effect of commutes for both samples. While we find that the influence of commute time on labor market outcomes varies by education level, the inclusion of this interaction does not eliminate differences in the effect of commutes across the supervision level samples. As in previous analyses, the primary effects of commute are identified for the high supervision occupation sample even after allowing these effect to vary by education level.

Table 8 and 9 provide the results for the employment and wage models, respectively.

In the baseline employment analysis, the negative effect of commute time for non-college educated workers in high supervision occupations is approximately the same magnitude as the effect for all high supervision occupation workers in Table 6, and the commute time effect for non-college educated workers in low supervision occupations is zero. The negative effect of commute time increases for college educated workers in both samples, but the education effect is significantly larger for high supervision occupation workers. As in Tables 4 and 6, the differences between the estimates for the high and low supervision occupation samples disappear as we add controls for neighborhood income and demographic composition, which appears to be consistent with a weak instrument problem that is biasing the IV estimates towards OLS estimates. In all of the wage analyses (Table 9), the effect of commute time for non-college educated workers in high supervision occupations is about half the magnitude of effect for all high supervision occupation workers in Table 6, and the effect of commute time for high supervision workers who are college educated is much larger, between two and a half and four times the estimated effect in Table 6. However, as before, there is no relationship between commute time and wages for workers in highly supervised occupations regardless of education level. While this analysis finds evidence that education matters in the relationship between commute time and labor market outcomes, the analysis still strongly supports the idea that this relationship arises primarily for workers in high supervision occupations especially in the case of the wage estimates.

[Insert Tables 8 and 9 here]

6 Summary and conclusions

This paper extends the classic model of efficiency wages to allow for substitution between shirking and leisure. This model is set within an urban equilibrium. The paper shows that a worker's benefit from shirking will depend upon their time endowment net of commuting time and that some workers will shirk in equilibrium unless firms can set efficiency wages based on each worker's commute. The model implies an empirical relationship between commutes and either employment and/or wages.

The paper uses data from the Public Use Microdata Sample of the 2000 U.S. Decennial Census to investigate whether the empirical implications of our efficiency wage-substitution model hold. The sample contains detailed information on residential and employment location, as well as standard demographic data including a worker's industry and occupation, labor market outcomes, and commuting patterns. A variety of models are estimated to examine whether a household's predicted commute time influences employment or labor market

earnings. The analyses find evidence to suggest that longer commutes lead to lower levels of employment among workers in highly supervised occupations presumably because these are occupations where shirking can have a large impact on firm productivity. The findings suggest that leisure and shirking are substitutable so that an increase in commutes reduces an individual's net time endowment leading to lower levels of leisure and higher levels of shirking.

The analysis also finds a strong positive relationship between wages and predicted commute for highly supervised workers. These results support the hypothesis that firms can wage discriminate based on commutes reducing the amount of shirking by workers with long commutes. However, the wage discrimination/spatially based efficiency wage is incomplete allowing for a relationship between commutes in employment for workers in heavily supervised occupation. It is notable that the employment and wage results for highly supervised workers are consistent in that long commutes lead to a higher tendency for workers to shirk, which predicts higher unemployment and/or the paying of higher wages to prevent shirking. For lightly supervised workers, no robust evidence exists of a relationship between commutes and either employment or wages.

The estimates from models that interact education level and commute time indicate that the effect of commutes is larger for college educated workers. As before, the influence of college education on the estimated effect of commute time occurs primarily for the sample of workers in high supervision occupations. For the high supervision sample, the influence of commute time on employment for college educated workers is approximately double the estimated effect for non-college educated workers, while the influence of commute time on employment for college educated workers is between four and six times the effect for non-college educated workers. As in the earlier analyses, the evidence that commute matters for workers in the low supervision sample is much weaker, especially in the wage models.

A substantial, growing literature exists on the operation of efficiency wage models in urban economies (see e.g. Zenou and Smith, 1995; Smith and Zenou, 1997; Zenou, 2002; Brueckner and Zenou 2003; Zenou, 2006), and prior to this paper no empirical evidence has been offered to suggest that efficiency wages are important in explaining differences in labor market outcomes over space. This paper offers a first attempt to test for the influence of efficiency wages on urban outcomes and finds strong evidence to support the relevance of this theory. Additional empirical work is needed investigating the role that efficiency wages or other models of unemployment can play in explaining spatial variation in employment and earnings.

In addition, the model and findings presented in this paper have broader implications

for the efficiency wage literature. Traditional efficiency wage models suggest that firms can prevent shirking by paying efficiency wages, but this paper suggests a mechanism by which workers will differ in their likelihood of shirking even if their underlying preferences are the same. Specifically, workers that have a lower net endowment of time prior to making a work effort decision, either due to a longer commute or non-work related personal obligations, such as a disabled spouse, child, or parent, are likely to differ in their propensity to shirk. Accordingly, if firms are unable to perfectly distinguish between workers in different circumstances, they will rationally set wages so that some workers shirk in equilibrium. The evidence found in this paper is consistent with this form of equilibrium shirking.

References

- [1] Allgulin, M. and T. Ellingsen (2002), "Monitoring and pay," *Journal of Labor Economics*, 20, 201-216.
- [2] Baxter, M. and U.J. Jerman (1999), "Household production and the excess sensitivity of consumption to current income," *American Economic Review*, 89, 902-920.
- [3] Bayer, P. and S.L. Ross (2006), "Identifying individual and group effects in the presence of sorting: A neighborhood effects application," NBER Working Paper No. 12211.
- [4] Bayer, P., Ross, S.L. and G. Topa (2005), "Place of work and place of residence: Informal hiring networks and labor market outcomes," NBER Working Paper No. 11019.
- [5] Becker, G. (1965), "A theory of the allocation of time," *Economic Journal*, 75, 493-517.
- [6] Bound, J., Jaeger, D.A., and R.M. Baker (1995), "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak," *Journal of the American Statistical Association*, 90, 443-450.
- [7] Brueckner, J.K. and Y. Zenou (2003), "Space and unemployment: The labor-market effects of spatial mismatch," *Journal of Labor Economics*, 21, 242-266.
- [8] Card, D. and J. Rothstein (2005), "Racial segregation and the black-white test score gap," NBER Working Paper No. 12078.
- [9] Chen, P. and P-A. Edin (2002), "Efficiency wages and inter-industry wage differentials: A comparison across methods of pay," *Review of Economics and Statistics*, 84, 617-631.

- [10] Cutler, D. and E. Glaeser (1997), “Are ghettos good or bad?” *Quarterly Journal of Economics*, 112, 827-872.
- [11] Deng, Y., S.L. Ross, and S. Wachter (2003), “Racial differences in homeownership: The effect of residential location,” *Regional Science and Urban Economics*, 33, 517-556.
- [12] Dickens, W.T. and L.F. Katz (1987), “Inter-industry wage differences and industry characteristics,” In *Unemployment and the Structure of Labor Markets*, K. Lang and J.S. Leonard (Eds.), New York: Basil Blackwell, pp. 48-89.
- [13] Durlauf, S. (2004), “Neighborhood effects,” In *Handbook of Regional and Urban Economics, Volume 4*, V. Henderson and J.F. Thisse (Eds.), Amsterdam: Elsevier Science.
- [14] Epple, D. (1987), “Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products,” *Journal of Political Economy*, 95, 59-80.
- [15] Evans, W., Oates, W. and R. Schwab (1992), “Measuring peer group effects: A study of teenage behavior,” *Journal of Political Economy*, 10, 966-991.
- [16] Fujita, M. (1989), *Urban Economic Theory*, Cambridge University Press, Cambridge.
- [17] Gibbons, R. and L. Katz (1992), “Does unmeasured ability explained inter-industry wage differentials,” *Review of Economic Studies*, 59, 515-535.
- [18] Glaeser, E.L. (1996), “Discussion on ‘Spatial effects upon employment outcomes: The case of New Jersey teenagers’ by O’Regan and Quigley,” *New England Economic Review: Federal Reserve Bank of Boston*, May/June, 58-64.
- [19] Horace, W.C. and R.L. Oaxaca, (2003), “Old wine in new bottles: A sequential estimation technique for the LPM,” IZA Discussion Paper #703.
- [20] Hoxby, C. (2000), “Does competition among public schools benefit students and taxpayers?,” *American Economic Review*, 90, 1209-1238.
- [21] Ihlanfeldt, K. R. and D. L. Sjoquist, (1998), “The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform,” *Housing Policy Debate*, 6, 849-892.
- [22] Katz, L.F., Kling, J., and J. Liebman (2001), “Moving to opportunity in Boston: Early results of a randomized mobility experiment,” *Quarterly Journal of Economics*, 116, 607-54.

- [23] Krueger, A. and L. Summers (1987), "Reflections on the inter-industry wage structure," in *Unemployment and the Structure of Labor Markets*, K. Lang and J.S. Leonard (Eds.), New York: Basil Blackwell, pp. 17-47.
- [24] Krueger, A. and L. Summers (1988), "Efficiency wages and the inter-industry wage structure," *Econometrica*, 56, 259-293.
- [25] Lazear, E.P. (2000), "Performance pay and productivity," *American Economic Review*, 90, 1346-1361.
- [26] Levenson, A. and C. Zoghi (2006), "The strength of occupational indicators as a proxy for skill," Unpublished manuscript, University of Southern California.
- [27] Madden, J.F. (1985), "Urban wage gradients: Empirical evidence," *Journal of Urban Economics*, 18, 291-301.
- [28] Manning, A. (2003), "The real thin theory: Monopsony in modern labour markets," *Labour Economics*, 10 105-131.
- [29] Moffitt, R. (2001), "Policy interventions, low-level equilibria and social interactions," In *Social Dynamics*, S.N. Durlauf and H.P. Young (Eds.), Cambridge, MA: MIT Press, pp. 45-82.
- [30] Murphy, K.M. and R.H. Topel (1987), "Unemployment, risk, and earnings: Testing for equalizing wage differences in the labor market," In *Unemployment and the Structure of Labor Markets*, K. Lang and J.S. Leonard (Eds.), New York: Basil Blackwell, pp. 103-140.
- [31] Murphy, K.M. and R.H. Topel (1990), "Efficiency wage reconsidered: Theory and evidence," In *Advances in the Theory and Measurement of Unemployment*, Y. Weiss and G. Fishelson (Eds.), New York: St. Martin's Press, pp. 104-240.
- [32] Neal, D. (1993), "Supervision and wages across industries," *Review of Economics and Statistics*, 75, 409-417.
- [33] O'Regan, K.M. and J.M. Quigley (1996), "Spatial effects upon employment outcomes: The case of New Jersey teenagers," *New England Economic Review*, May-June 1996.
- [34] Paarsh, H.J. and B. Shearer (2000), "Piece rates, fixed wages, and incentive effects: Statistical evidence from payroll records," *International Economic Review*, 41, 59-92.

- [35] Ross, S.L. (1998), "Racial differences in residential and job mobility: Evidence concerning the spatial mismatch hypothesis," *Journal of Urban Economics*, 43, 112-135.
- [36] Ross, S.L. and J. Yinger, (1995), "A comparative static analysis of an open urban model with a full labor market and suburban employment," *Regional Science and Urban Economics*, 25, 575-605.
- [37] Ross, S.L. and Y. Zenou (2007), "Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory," CEPR Discussion Paper No. 6128.
- [38] Shapiro, C. and J.E. Stiglitz (1984), "Equilibrium unemployment as a worker discipline device," *American Economic Review*, 74, 433-444.
- [39] Smith, T.E. and Y. Zenou (1997), "Dual labor markets, urban unemployment, and multicentric cities," *Journal of Economic Theory*, 76, 185-214.
- [40] Timothy, D. and W.C. Wheaton (2001), "Intra-urban wage variation, employment location, and commuting time," *Journal of Urban Economics*, 50, 338-336.
- [41] Topa, G. (2001), "Social interactions, local spillovers, and unemployment," *Review of Economic Studies*, Vol. 68, pp. 261-295.
- [42] Weinberg, B.A. (2000), "Black residential centralization and the spatial mismatch hypothesis," *Journal of Urban Economics*, 48, 110-134.
- [43] Weinberg, B.A., Reagan, P.B. and J.J. Yankow (2005), "Do neighborhoods affect hours worked: Evidence from longitudinal data," *Journal of Labor Economics*, 22, 891-924.
- [44] White, M.J. (1999), "Urban models with decentralized employment: Theory and empirical work," In *Handbook of Regional and Urban Economics Vol. 3*, P. Cheshire and E.S. Mills (Eds.), Amsterdam: Elsevier Science, pp. 1375-1412.
- [45] Zax J.S. (1991), "Compensation for commutes in labor and housing markets," *Journal of Urban Economics*, 30, 192-207.
- [46] Zax, J.S. and J.F. Kain (1991), "Commutes, quits and moves," *Journal of Urban Economics*, 29, 153-165.
- [47] Zenou, Y. (2002), "How do firms redline workers?" *Journal of Urban Economics*, 52, 391-408.

- [48] Zenou, Y. (2006), “Efficiency wages and unemployment in cities: The case of high relocation costs,” *Regional Science and Urban Economics*, 36, 49-71.
- [49] Zenou, Y. (2008), *Urban Labor Economics*, Cambridge: Cambridge University Press, forthcoming.
- [50] Zenou, Y. and T.E. Smith (1995), “Efficiency wages, involuntary unemployment and urban spatial structure,” *Regional Science and Urban Economics*, 25, 821-845.
- [51] Zivot, E., Startz, R., and C.R. Nelson (1998), “Valid confidence intervals and inference in the presence of weak instruments,” *International Economic Review*, 39, 1119-1144.

APPENDIX

Proof of Proposition 1

First observe that

$$-\frac{\partial \Psi^{NS}(x, \bar{I}_{no})}{\partial x} \Big|_{x=\tilde{x}_{no}} \geq -\frac{\partial \Psi^S(x, \bar{I}_{no})}{\partial x} \Big|_{x=\tilde{x}_{no}}$$

This is equivalent to:

$$\tau + t(1 - u^{NS}) \frac{\partial V(1 - T - tx, \bar{e})}{\partial l_{no}} \Big|_{x=\tilde{x}_{no}} \geq \tau + t(1 - u^S) \frac{\partial V(1 - T - tx, \underline{e})}{\partial l_{no}} \Big|_{x=\tilde{x}_{no}}$$

or

$$(1 - u^{NS}) \frac{\partial V(1 - T - t\tilde{x}_{no}, \bar{e})}{\partial l_{no}} \Big|_{x=\tilde{x}_{no}} \geq (1 - u^S) \frac{\partial V(1 - T - tx, \underline{e})}{\partial l_{no}} \Big|_{x=\tilde{x}_{no}}$$

Now, using assumption (2), i.e.

$$\frac{\partial^2 V(l, e)}{\partial l \partial e} > 0$$

then it is easy to verify that

$$(1 - u^{NS}) \frac{\partial V(1 - T - t\tilde{x}_{no}, \bar{e})}{\partial l_{no}} \Big|_{x=\tilde{x}_{no}} > (1 - u^S) \frac{\partial V(1 - T - tx, \underline{e})}{\partial l_{no}} \Big|_{x=\tilde{x}_{no}}$$

is always true. ■

Proof of Lemma 1. Differentiation of \tilde{x}_{no} yields:

$$\frac{\partial^2 \tilde{x}_{no}}{\partial w_{no}^2} = \frac{T(u^S - u^{NS})}{t \left[(1 - u^{NS}) \frac{\partial \tilde{V}_{no}^{NS}}{\partial l_{no}} - (1 - u^S) \frac{\partial \tilde{V}_{no}^S}{\partial l_{no}} \right]^2} \left[(1 - u^{NS}) \frac{\partial^2 \tilde{V}_{no}^{NS}}{\partial l_{no}^2} - (1 - u^S) \frac{\partial^2 \tilde{V}_{no}^S}{\partial l_{no}^2} \right] \frac{\partial \tilde{x}_{no}}{\partial w_{no}}$$

Thus, if (14) holds and since $1 - u^{NS} > 1 - u^S$, we have:

$$(1 - u^{NS}) \frac{\partial^2 \tilde{V}_{no}^{NS}}{\partial l_{no}^2} - (1 - u^S) \frac{\partial^2 \tilde{V}_{no}^S}{\partial l_{no}^2} < 0$$

As a result,

$$\text{sgn} \frac{\partial^2 \tilde{x}_{no}}{\partial w_{no}^2} = \text{sgn} \frac{\partial \tilde{x}_{no}}{\partial w_{no}}$$

Using Proposition 2, the result is straightforward. ■

Proof of Proposition 3

First order conditions for profit maximization yield:

$$\begin{aligned} \frac{\partial \Pi_{no}}{\partial w_{no}} &= F'(\cdot) \frac{\partial \tilde{x}_{no}}{\partial w_{no}} [(1 - u^{NS})\bar{e} - (1 - u^S)\underline{e}] - [(1 - \tilde{x}_{no})(1 - u^S) + \tilde{x}_{no}(1 - u^{NS})] \quad (28) \\ &\quad - w_{no} \frac{\partial \tilde{x}_{no}}{\partial w_{no}} (u^S - u^{NS}) = 0 \end{aligned}$$

$$\frac{\partial \Pi_{no}}{\partial \alpha_{no}} = F'(\cdot) [(1 - \tilde{x}_{no})(1 - u^S)\bar{e} + \tilde{x}_{no}(1 - u^{NS})\underline{e}] - w_{no} [(1 - \tilde{x}_{no})(1 - u^S) + \tilde{x}_{no}(1 - u^{NS})] = 0 \quad (29)$$

Now by combining these two equations, we obtain the following equation that determines the wage setting:

$$w_{no} \frac{\partial \tilde{x}_{no}}{\partial w_{no}} = \frac{L_{no} [(1 - \tilde{x}_{no})(1 - u^S)\bar{e} + \tilde{x}_{no}(1 - u^{NS})\underline{e}]}{(\bar{e} - \underline{e}) [(1 - \tilde{x}_{no})(1 - u^S)^2 + \tilde{x}_{no}(1 - u^{NS})^2]} \quad (30)$$

where $L_{no} = (1 - \tilde{x}_{no})(1 - u^S) + \tilde{x}_{no}(1 - u^{NS})$, whereas the employment level in each firm is determined by (29).

We would like now to show that $\partial^2 \Pi_{no} / \partial w_{no}^2 < 0$. By Differentiating (28) with respect to w_{no} , we easily obtain:

$$\begin{aligned} \frac{\partial^2 \Pi_{no}}{\partial w_{no}^2} &= \alpha_{no} \frac{\partial^2 \tilde{x}_{no}}{\partial w_{no}^2} \{ F'(\cdot) [(1 - u^{NS})\bar{e} - (1 - u^S)\underline{e}] - w_{no}(u^S - u^{NS}) \} \\ &\quad + \alpha_{no} \frac{\partial \tilde{x}_{no}}{\partial w_{no}} \left\{ \alpha_{no} \frac{\partial \tilde{x}_{no}}{\partial w_{no}} F''(\cdot) [(1 - u^{NS})\bar{e} - (1 - u^S)\underline{e}]^2 - (u^S - u^{NS}) \right\} \end{aligned}$$

Assume that (2) holds. Then, from Proposition 2, we know that $\partial \tilde{x}_{no} / \partial w_{no} > 0$. Using the fact that $F''(\cdot) < 0$ and $u^S > u^{NS}$, we obtain:

$$\alpha_{no} \frac{\partial \tilde{x}_{no}}{\partial w_{no}} \left\{ \alpha_{no} \frac{\partial \tilde{x}_{no}}{\partial w_{no}} F''(\cdot) [(1 - u^{NS})\bar{e} - (1 - u^S)\underline{e}]^2 - (u^S - u^{NS}) \right\} < 0$$

Now, (28) can be written as:

$$\frac{\partial \tilde{x}_{no}}{\partial w_{no}} [F'(\cdot) [(1 - u^{NS})\bar{e} - (1 - u^S)\underline{e}] - w_{no}(u^S - u^{NS})] = L_{no}$$

Since $\partial \tilde{x}_{no} / \partial w_{no} > 0$, this implies that

$$F'(\cdot) [(1 - u^{NS})\bar{e} - (1 - u^S)\underline{e}] - w_{no}(u^S - u^{NS}) > 0$$

Finally, using Lemma 1 and assuming (14) imply that $\partial^2 \tilde{x}_{no} / \partial w_{no}^2 < 0$ and thus

$$\alpha_{no} \frac{\partial^2 \tilde{x}_{no}}{\partial w_{no}^2} \{F'(\cdot) [(1 - u^{NS})\bar{e} - (1 - u^S)\underline{e}] - w_{no}(u^S - u^{NS})\} < 0$$

Consequently, $\partial^2 \Pi_{no} / \partial w_{no}^2 < 0$. ■

Definition of equilibrium when firms do not observe workers' location

Definition 1 Consider the case when firms do not observe workers' location. Assume (2) and (14). Then, the equilibrium $(\bar{I}_{no}, \tilde{x}_{no}, w_{no}, \alpha_{no}, M_{no}, u^{NS}, u^S)$ such that

$$\Psi^S(1, \bar{I}_{no}) = 0 \tag{31}$$

(12), (28), (29), (17), (8), and (9) hold.

The conditions of the land market equilibrium are given by (12) and (31). These conditions guarantee that the equilibrium land rent has to be continuous over all the city, i.e. land rents of shirkers and nonshirkers have to be equal at the intersection location \tilde{x} and the land rent at the city fringe has to be equal to the agricultural land rent (here normalized to zero). Conditions of the labor market equilibrium are given by the five other equations. An equilibrium requires solving *simultaneously* these two equilibria.

Definition of equilibrium when firms do observe workers' location

We focus on a symmetric labor market equilibrium in which each firm employs the same number of workers $\alpha_o L_o = L_o / M_o$ so that

$$M_o = \frac{1}{\alpha_o} \tag{32}$$

Definition 2 Consider the case when firms can wage discriminate in terms of location. Assume (22). Then, the equilibrium is a vector $(w_o(x), \alpha_o, M_o, u^{NS}, u^S, \bar{I}_o)$ such that (21), (24), (29), (32), (8), (9) and (20) hold.

Table 1: Means and Standard Errors		
Variable Names	Occupations with Supervision	Occupations without Supervision
Employment	0.954 (0.21)	0.981 (0.14)
Adult's Housing Wage	17.480 (15.71)	29.531 (26.42)
Years in Labor Market	2.009 (0.86)	18.322 (8.585)
Adult has Masters Degree or More	0.031 (0.17)	0.300 (0.46)
Adult is College Graduate	0.130 (0.34)	0.377 (0.48)
Adult has Two Years of College	0.069 (0.25)	0.070 (0.25)
Adult is High School Graduate	0.575 (0.49)	0.233 (0.42)
Adult is Hispanic	0.178 (0.38)	0.055 (0.23)
Adult is African-American	0.126 (0.33)	0.068 (0.25)
Adult is Asian or Pacific Islander	0.049 (0.22)	0.084 (0.28)
Adult is Native American	0.005 (0.07)	0.003 (0.05)
Adult is Another Race	0.120 (0.33)	0.040 (0.20)
Household Contains Children	0.503 (0.50)	0.520 (0.50)
Adult is Married	0.624 (0.48)	0.699 (0.46)
Non-Native Adult Residing <= 10 years	0.084 (0.28)	0.062 (0.24)
Non-Native Adult Residing > 10 years	0.158 (0.36)	0.118 (0.32)
Commuting Time in Hours	0.492 (0.349)	0.512 (0.347)
Standardized Employment Access	-0.053 (0.989)	0.083 (1.021)
Mean Family Income in \$1,000's	78.148 (24.376)	87.295 (26.276)
Fraction African-American	0.133 (0.178)	0.104 (0.142)
Fraction Hispanic	0.149 (0.178)	0.112 (0.127)
Fraction Non-Naïve	0.225 (0.192)	0.213 (0.159)
Sample Size	706,296	421,311

Notes: The sample includes all males aged 25 to 55 who are currently participating in the labor market and not residing in group quarters. Standard errors are shown in parentheses. The first column is based on all individuals who report an occupation which belongs to a general occupation categories classified as having high levels of supervision. The next column is based on a sample containing individuals in all other occupations. Average wage is based on a smaller sample restricted to individuals working at least 40 weeks, an average of at least 30 hours per week, and with estimated hourly earnings at the Federal minimum wage or higher. The wage rate sample size for occupations with supervision is 545,727 and for occupations without supervision is 357,528. Commute times are only reported for individuals who are employed at the time of the survey. The respective sample sizes for the commute time mean and standard deviation are 648,816 and 394,391.

Variable Names	Occupations with Supervision		Occupations without Supervision	
	OLS	IV	OLS	IV
Years in Labor Market / 10	-0.002 (1.19)	-0.002 (1.24)	-0.010 (6.85)	-0.010 (6.85)
Square Years in Labor Market / 10	0.001 (1.97)	0.001 (2.04)	0.001 (3.09)	0.001 (3.09)
Adult has Masters Degree or More	0.033 (17.13)	0.033 (16.92)	0.034 (9.09)	0.034 (9.09)
Adult is College Graduate	0.043 (24.55)	0.043 (24.10)	0.029 (8.35)	0.029 (8.36)
Adult has Two Years of College	0.037 (21.04)	0.038 (20.98)	0.026 (7.38)	0.026 (7.38)
Adult is High School Graduate	0.027 (19.38)	0.027 (19.30)	0.021 (6.93)	0.021 (6.93)
Adult is Hispanic	-0.002 (1.12)	-0.001 (1.01)	-0.005 (2.51)	-0.005 (2.51)
Adult is African-American	-0.047 (23.80)	-0.047 (23.64)	-0.016 (13.12)	-0.016 (13.10)
Adult is Asian or Pacific Islander	-0.015 (8.83)	-0.015 (8.76)	-0.004 (2.97)	-0.004 (2.98)
Adult is Native American	-0.029 (5.68)	-0.029 (5.70)	-0.012 (2.28)	-0.012 (2.28)
Adult is Another Race	-0.005 (3.82)	-0.005 (3.77)	-0.002 (1.28)	-0.002 (1.29)
Household Contains Children	0.009 (13.28)	0.009 (13.34)	0.007 (10.51)	0.007 (10.56)
Adult is Married	0.026 (28.33)	0.026 (28.21)	0.012 (16.91)	0.012 (16.90)
Non-Native Residing <= 10 years	0.017 (9.06)	0.017 (9.10)	-0.004 (4.04)	-0.004 (4.01)
Non-Native Residing > 10 years	0.008 (6.03)	0.008 (6.13)	-0.002 (2.43)	-0.002 (2.44)
Commute Time		-0.024 (4.06)		-0.002 (0.42)
Sample Size	653,624	650,133	392,917	391,966

Notes: The models are estimated using a sample of labor force participants and a linear probability model iteratively in order to eliminate observations that predict residuals outside of the zero-one interval. All models control for industry and Occupation-MSA fixed effects, as well as controls for the housing attributes that are used to predict exposure to residential location attributes. All standard errors are clustered at the Occupation-MSA level in order to address heteroscedasticity that might arise in the fixed effects specification. T-statistics are shown in parentheses.

Variable Names	Occupations with Supervision		Occupations without Supervision	
	OLS	IV	OLS	IV
Years in Labor Market / 10	0.202 (31.50)	0.202 (31.46)	0.349 (34.83)	0.349 (34.82)
Square Years in Labor Market / 10	-0.031 (17.96)	-0.031 (17.92)	-0.055 (21.37)	-0.055 (21.37)
Adult has Masters Degree or More	0.471 (25.05)	0.471 (25.10)	0.565 (37.52)	0.565 (37.56)
Adult is College Graduate	0.374 (34.82)	0.374 (34.82)	0.429 (31.15)	0.429 (31.16)
Adult has Two Years of College	0.199 (25.19)	0.199 (25.22)	0.197 (16.06)	0.197 (16.05)
Adult is High School Graduate	0.136 (22.54)	0.136 (22.55)	0.154 (13.88)	0.154 (13.89)
Adult is Hispanic	-0.091 (12.91)	-0.092 (12.96)	-0.088 (8.73)	-0.088 (8.73)
Adult is African-American	-0.094 (15.63)	-0.094 (15.66)	-0.142 (16.19)	-0.142 (16.20)
Adult is Asian or Pacific Islander	-0.130 (10.24)	-0.131 (10.25)	-0.060 (3.80)	-0.060 (3.81)
Adult is Native American	-0.086 (8.89)	-0.086 (8.87)	-0.124 (8.48)	-0.124 (8.48)
Adult is Another Race	-0.046 (7.64)	-0.046 (7.65)	-0.086 (9.43)	-0.086 (9.45)
Household Contains Children	0.043 (20.20)	0.043 (20.17)	0.013 (3.55)	0.013 (3.57)
Adult is Married	0.083 (30.11)	0.083 (30.16)	0.057 (9.76)	0.058 (9.79)
Non-Native Residing <= 10 years	-0.186 (21.04)	-0.186 (21.04)	0.003 (0.25)	0.003 (0.25)
Non-Native Residing > 10 years	-0.084 (15.08)	-0.084 (15.18)	-0.028 (3.72)	-0.028 (3.71)
Commute Time		0.082 (3.70)		-0.052 (1.99)
R-Square	0.318	0.318	0.293	0.293
Sample Size	545,727		357,528	

Notes: The models are estimated by applying two stage least squares to a sample of individuals working at least 40 weeks, an average of at least 30 hours per week, and with estimated hourly earnings at the Federal minimum wage or higher. All models control for industry and Occupation-MSA fixed effects, as well as housing attribute controls. All standard errors are clustered at the Occupation-MSA level, and T-statistics are shown in parentheses.

Table 4: Parameter Estimates and T-Statistics for Employment Models ¹				
Variable Names	Baseline Model	Employment Access	Plus Average Income	Plus Demographic Controls
Occupations with Supervision				
Commute Time	-0.024 (4.06)	-0.023 (3.48)	-0.011 (1.96)	-0.005 (0.87)
Employment Access		0.001 (0.38)	0.000 (0.05)	0.004 (2.39)
Mean Family Income / \$100,000			0.053 (7.08)	0.026 (3.27)
Fraction African-American				-0.062 (5.55)
Fraction Hispanic				-0.091 (4.89)
Fraction Non-Native				0.062 (3.59)
Sample Size	650,133	650,367	649,898	641,599
Occupations without Supervision				
Commute Time	-0.002 (0.42)	0.001 (0.17)	-0.001 (0.13)	0.002 (0.38)
Employment Access		0.001 (1.16)	0.000 (0.40)	0.001 (0.45)
Mean Family Income / \$100,000			0.018 (2.99)	0.008 (1.22)
Fraction African-American				-0.009 (0.92)
Fraction Hispanic				0.006 (0.37)
Fraction Non-Native				-0.034 (1.93)
Sample Size	391,966	392,004	392,149	383,328

Notes: The models are estimated using an iterative linear probability model. All models control for the demographic attributes shown in Tables 2 and 3 plus housing attributes, industry and Occupation-MSA fixed effects. All standard errors are clustered at the Occupation-MSA level, and T-statistics are shown in parentheses.

Table 5: Parameter Estimates and T-Statistics for Log Wage Models ¹				
Variable Names	Baseline Model	Employment Access	Plus Average Income	Plus Demographic Controls
Occupations with Supervision				
Commute Time	0.082 (3.70)	0.109 (3.07)	0.137 (3.70)	0.122 (3.00)
Employment Access		0.016 (1.15)	0.013 (0.91)	0.014 (1.14)
Mean Family Income / \$100,000			0.140 (3.65)	0.189 (4.81)
Fraction African-American				0.069 (1.53)
Fraction Hispanic				-0.144 (1.06)
Fraction Non-Native				0.209 (1.50)
R-Square	0.318	0.318	0.318	0.318
Sample Size	545,727			
Occupations without Supervision				
Commute Time	-0.052 (1.99)	0.014 (0.42)	0.005 (0.15)	-0.006 (0.17)
Employment Access		0.030 (2.49)	0.022 (1.98)	0.023 (2.03)
Mean Family Income / \$100,000			0.142 (2.18)	0.202 (3.27)
Fraction African-American				-0.091 (0.94)
Fraction Hispanic				0.272 (1.32)
Fraction Non-Native				0.043 (0.25)
R-Square	0.293	0.293	0.293	0.294
Sample Size	357,528			

Notes: The models are estimated using two stage least squares. All models control for the demographic attributes shown in Tables 2 and 3 plus housing attributes, industry and Occupation-MSA fixed effects. All standard errors are clustered at the Occupation-MSA level, and T-statistics are shown in parentheses.

Table 6: Employment Models with Smaller Set of Instruments ¹				
Variable Names	Baseline Model	Employment Access	Plus Average Income	Plus Demographic Controls
Occupations with Supervision				
Commute Time	-0.027 (4.57)	-0.025 (3.68)	-0.014 (2.39)	-0.008 (1.43)
Employment Access		0.001 (0.79)	0.001 (0.46)	0.005 (3.13)
Mean Family Income / \$100,000			0.045 (5.76)	0.020 (2.37)
Fraction African-American				-0.055 (4.72)
Fraction Hispanic				-0.100 (5.34)
Fraction Non-Native				0.071 (4.09)
Sample Size	653,086	653,361	652,838	644,995
Occupations without Supervision				
Commute Time	-0.002 (0.46)	0.002 (0.33)	0.000 (0.07)	0.003 (0.57)
Employment Access		0.001 (1.25)	0.001 (0.69)	0.001 (0.71)
Mean Family Income / \$100,000			0.015 (2.20)	0.004 (0.60)
Fraction African-American				-0.008 (0.79)
Fraction Hispanic				0.003 (0.13)
Fraction Non-Native				-0.032 (1.69)
Sample Size	395,273	395,390	395,623	384,131

Notes: The models are estimated using two stage least squares where the set of housing attribute instruments have been reduced by eliminating possibly endogenous housing attributes like no kitchen in unit or mobile home. All models control for the demographic attributes shown in Tables 2 and 3 plus housing attributes, industry and Occupation-MSA fixed effects. All standard errors are clustered at the Occupation-MSA level, and T-statistics are shown in parentheses.

Table 7: Log Wage Models with Smaller Set of Instruments ¹				
Variable Names	Baseline Model	Employment Access	Plus Average Income	Plus Demographic Controls
Occupations with Supervision				
Commute Time	0.067 (2.80)	0.092 (2.15)	0.117 (2.68)	0.094 (1.96)
Employment Access		0.013 (0.85)	0.010 (0.67)	0.011 (0.82)
Mean Family Income / \$100,000			0.115 (3.05)	0.187 (4.68)
Fraction African-American				0.091 (1.97)
Fraction Hispanic				-0.143 (1.06)
Fraction Non-Native				0.252 (1.80)
R-Square	0.314	0.314	0.314	0.314
Sample Size	355,727			
Occupations without Supervision				
Commute Time	-0.052 (2.00)	0.022 (0.60)	0.014 (0.40)	-0.003 (0.07)
Employment Access		0.030 (2.32)	0.024 (1.91)	0.024 (1.97)
Mean Family Income / \$100,000			0.120 (1.87)	0.189 (3.03)
Fraction African-American				-0.084 (0.82)
Fraction Hispanic				0.230 (1.08)
Fraction Non-Native				0.099 (0.57)
R-Square	0.290	0.290	0.290	0.290
Sample Size	357,528			

Notes: The models are estimated using two stage least squares where the set of housing attribute instruments have been reduced by eliminating possibly endogenous housing attributes like no kitchen in unit or mobile home. All models control for the demographic attributes shown in Tables 2 and 3 plus housing attributes, industry and Occupation-MSA fixed effects. All standard errors are clustered at the Occupation-MSA level, and T-statistics are shown in parentheses.

Table 8: Employment Models with Education Interaction ¹				
Variable Names	Baseline Model	Employment Access	Plus Average Income	Plus Demographic Controls
Occupations with Supervision				
Commute Time	-0.025 (4.12)	-0.023 (3.28)	-0.012 (2.07)	-0.007 (1.12)
Commute Time*Four Year Degree	-0.022 (4.22)	-0.021 (4.18)	-0.015 (2.65)	-0.016 (3.12)
Employment Access		0.001 (0.80)	0.001 (0.52)	0.005 (3.12)
Mean Family Income / \$100,000			0.045 (5.74)	0.019 (2.32)
Fraction African-American				-0.055 (4.73)
Fraction Hispanic				-0.098 (5.25)
Fraction Non-Native				0.070 (4.05)
Sample Size	654,152	654,356	653,051	647,268
Occupations without Supervision				
Commute Time	0.005 (0.83)	0.008 (1.30)	0.006 (1.07)	0.010 (1.54)
Commute Time*Four Year Degree	-0.011 (2.34)	-0.010 (2.27)	-0.010 (2.23)	-0.012 (2.59)
Employment Access		0.001 (1.23)	0.001 (0.54)	0.001 (0.50)
Mean Family Income / \$100,000			0.015 (2.18)	0.005 (0.71)
Fraction African-American				-0.007 (0.74)
Fraction Hispanic				0.002 (0.11)
Fraction Non-Native				-0.032 (1.65)
Sample Size	395,602	396,766	398,321	385,678

Notes: The models that include an interaction between having a four year college degree and commute time are estimated using two stage least squares where the set of housing attribute instruments have been reduced by eliminating possibly endogenous housing attributes like no kitchen in unit or mobile home. All models control for the demographic attributes shown in Tables 2 and 3 plus housing attributes, industry and Occupation-MSA fixed effects. All standard errors are clustered at the Occupation-MSA level, and T-statistics are shown in parentheses.

Table 9: Log Wage Models with Education Interaction ¹				
Variable Names	Baseline Model	Employment Access	Plus Average Income	Plus Demographic Controls
Occupations with Supervision				
Commute Time	0.028 (1.16)	0.049 (1.82)	0.074 (2.71)	0.048 (1.68)
Commute Time*Four Year Degree	0.224 (2.26)	0.222 (2.29)	0.222 (2.27)	0.224 (2.30)
Employment Access		0.011 (0.79)	0.008 (0.58)	0.009 (0.74)
Mean Family Income / \$100,000			0.114 (3.05)	0.190 (4.83)
Fraction African-American				0.094 (2.07)
Fraction Hispanic				-0.147 (1.13)
Fraction Non-Native				0.263 (2.02)
R-Square	0.314	0.314	0.314	0.314
Sample Size	545,727			
Occupations without Supervision				
Commute Time	-0.066 (1.99)	0.008 (0.21)	0.001 (0.03)	-0.016 (0.45)
Commute Time*Four Year Degree	0.021 (0.65)	0.021 (0.64)	0.020 (0.61)	0.021 (0.65)
Employment Access		0.030 (2.32)	0.024 (1.91)	0.024 (1.97)
Mean Family Income / \$100,000			0.119 (1.88)	0.189 (3.03)
Fraction African-American				-0.083 (0.81)
Fraction Hispanic				0.228 (1.08)
Fraction Non-Native				0.103 (0.60)
R-Square	0.290	0.290	0.290	0.290
Sample Size	357,528			

Notes: The models that include an interaction between having a four year college degree and commute time are estimated using two stage least squares where the set of housing attribute instruments have been reduced by eliminating possibly endogenous housing attributes like no kitchen in unit or mobile home. All models control for the demographic attributes shown in Tables 2 and 3 plus housing attributes, industry and Occupation-MSA fixed effects. All standard errors are clustered at the Occupation-MSA level, and T-statistics are shown in parentheses.