

Spill-Overs from Good Jobs: A New Approach to a Recurring Debate

Paul Beaudry, David A. Green and Benjamin Sand *

February 2007

Abstract

Does attracting or losing jobs in high paying sectors have important spill-over effects on wages in other sectors? The answer to this question is central to a proper assessment of many trade and industrial policies. In this paper, we exploit a search and matching model of the labor market to clarify how this question can be properly posed and how it can be empirically explored. In our empirical implementation, we use U.S. Census data over the years 1970 to 2000 to quantify the relationship between changes in industrial composition and changes in industry-specific city-level wages. Our findings are that sectoral level wages act as strategic complements in a manner consistent with bargaining theory, and that the spill-over (i.e., general equilibrium) effects associated with changes in the fraction of jobs in high paying sectors are very substantial and persistent. Our point estimates indicate that the total effect on average wages of a change in industrial composition that favors high paying sectors is about 3.5 times that obtained from a commonly used composition-adjustment approach which neglects general equilibrium effects. We interpret our results as highlighting the relevance of search and matching models for understanding wage determination in a decentralized economy, and we argue that our results provide considerable support to the populist view that changes in industrial composition are very important for understanding labor market outcomes.

Key Words: Wages, Industrial Composition, Search, Bargaining

JEL Class.: E32

Paul Beaudry
Department of Economics
University of British Columbia
997-1873 East Mall
Vancouver, B.C.
Canada, V6T 1Z1
and NBER.
paulbe@interchange.ubc.ca

David A. Green
Department of Economics
University of British Columbia
997-1873 East Mall
Vancouver, B.C.
Canada, V6T 1Z1
Research Associate, IFS, London
green@econ.ubc.ca

Benjamin Sand
Department of Economics
University of British Columbia
997-1873 East Mall
Vancouver, B.C.
Canada, V6T 1Z1
bmsand@interchange.ubc.ca

*The authors thank M. Bombardini, D. Card, G. Dahl, M. Doms, J. Fernald, N. Fortin, G. Galipoli, R. Gordon, M. Greenstone, S. Kortum, I. King, T. Lemieux, K. Milligan and B. Meyer for helpful discussions.

Introduction

In popular discussion about labor market developments, whether it be at the local or national level, changes in the nature of jobs are often given a pre-eminent role. In particular, it is often claimed that labor market performance hinges on whether an economy is attracting or losing “good jobs”: that is, jobs in industries which pay a premium relative to wages for similarly qualified workers in other industries. For example, in Bluestone and Harrison’s highly cited 1982 book, *The Deindustrialization of America*, the authors argued that the loss of highly paid manufacturing jobs was key to understanding the poor labor market performance of the U.S. economy during the 1970s and 1980s. However, based on simple accounting exercises aimed at assessing the potential importance of changes in industrial composition on average wages, most serious economic researchers have dismissed such views as being ill-informed. Interestingly, the populist view is now reemerging in economic discussions in many countries because of perceptions about the effects of globalization on industrial composition and wages. For this reason it appears to be an opportune time to re-visit the issue of whether the populist view holds any validity.

Debates on the effects of attracting “good jobs” reduce, in the main, to the question of whether and how wages are affected by a shift in employment from one sector to another, holding total employment constant. From a research perspective, attempts to answer this question raise a further series of questions. Is the question properly posed in the first place? How can it be addressed given that industrial composition is endogenous? Can an answer be inferred from non-experimental data? The object of this paper is to shed new light on these questions. In particular, we show that a search and matching model of the labor market offers a coherent framework within which the question can be posed, and, further, it provides clear guidance on how it can be addressed empirically. As we will show, with multiple sectors and search, wages in different sectors interact as strategic complements. This strategic complementarity is then a potential source of general equilibrium (spill-over) effects from shifts in industrial composition to wages.

It is worthwhile contrasting our approach with the common practice of using a simple accounting procedure for examining the effects of a shift in industrial composition on average wages. The accounting approach (also known as the shift-share analysis) consists of multiplying the change in the proportion of workers in a given industry in a city by the average wage in that industry then summing across industries. The result from such an exercise almost always indicates that the average wage change directly accounted for by changes in sectoral composition of employment is small. The validity of the accounting approach hinges

critically on the assumption that a change in employment composition between two sectors does not affect wages in other sectors if the change does not effect the employment level in these other sectors; i.e., there are no general equilibrium effects on wages of a shift in the composition of the industrial structure. Without the assumption, one would also have to account for changes in the within sector wages arising from the compositional shifts, destroying the clean break into “within” (premia change) and “between” (composition change) components that is a key feature of the accounting approach.¹

Our goal in this paper is to examine whether changes in sectoral composition of employment, especially shifts in composition between high and low paying sectors, have important general equilibrium effects on the determination of within sector wages. To help clarify our question and to help direct the empirical strategy to answer it, we begin by showing how a standard search and matching model of the labor market,² extended to include many industries, implies spill-over effects of changes in industrial composition that would invalidate the use of the standard accounting approach. In particular, the search and matching framework implies that a change in industrial composition, through its effect on the bargaining position of workers, can affect wages in sectors not directly involved in the compositional change as long as workers are potentially mobile across sectors. In this case, an improved outside option for workers will place upward pressure on wages even if total employment is unaffected. This implication of search and matching is very basic and implies that wages in different sectors act as strategic complements. Although the insight is straightforward, to the best of our knowledge it is an empirical implication of this class of models which has not previously been extensively pursued even though it has important ramifications for our understanding of the functioning of the labor market.³ One of the reasons it may not have been extensively pursued is that its empirical exploration involves many hurdles, including solving a reflection problem and controlling for the endogeneity of industrial composition. The theoretical framework we use is particularly helpful in providing guidance on how to discuss and address issues related to endogeneity, potential instrumental variables, and interpretation.

¹ There are many ways to justify the no-GE effects assumption, which is part of the appeal of this approach. The easiest defense is to note that if wages are simply a function of productivity and returns to labor are close to constant, one just needs to assume that changes in industrial composition do not change productivity within sectors to arrive at the conclusion that there are no GE effects. The latter assumption might be viewed as innocuous by many economists, but it is the one we want to place into question.

² For an introduction to search and matching models of the labor market see Mortensen and Pissarides [1999]

³ One side implication of our model is a relationship between wages and the employment rate in a local economy that echoes that explored in Blanchflower and Oswald [1995]. We differ from that work, and from other research on bargaining effects on wage setting, in that we emphasize the importance of industrial composition on setting wages within all sectors in an economy.

In order to examine the relevance of spill-over effects associated with changes in the fraction of jobs that are “good jobs”, we exploit geographical variation in industrial composition across U.S. cities over the period 1970-2000. Our approach is to look at whether wage changes within a given industry vary systematically across cities with changes in the predicted distribution of employment across other industries in each city. While we do not impose the structure of our model on the data, our empirical specification is closely aligned with the theory we present. In particular, we examine whether wages in any given industry-city cell tend to increase in a city which has witnessed an (externally driven) change in industrial composition that is more concentrated toward high paying jobs. To do this, we examine 10 year changes in industry \times city level wages using data from the 1970, 1980, 1990 and 2000 U.S. Censuses for 152 cities.⁴ We devote considerable effort toward addressing endogeneity issues associated with a reflection problem inherently related to strategic complementarities, and we also address possible non-random selection in unobservable worker characteristics across cities using the method proposed in Dahl [2002].

The main empirical result of the paper is that city level changes in industrial composition have effects on average wages that are 3 to 3.5 times that implied by a pure accounting approach, indicating the presence of substantial general equilibrium effects. It is important to keep in mind when considering this result that measured composition effects are often small, so that the effects we find are large but not extreme.⁵ Nonetheless impacts of the size we identify have the effect of giving credence to explanations for changes in wage patterns that operate through changes in industrial structure; explanations which have largely been discounted because the pure accounting measures of their impacts are relatively small (e.g., Bound and Johnson [1992]). Furthermore, our results provide considerable support for search and matching models of the labor market as a framework for understanding wage determination in a decentralized environment. For example, we find that wages in different sectors act like strong strategic complements in precisely the manner predicted by decen-

⁴ There are both advantages and disadvantages of using city level observations to examine this issue. On the one hand, an attractive feature is there are over 150 metropolitan areas in the U.S. which gives us a sample with many different patterns of industrial composition. On the other, if labor markets across U.S. cities are all perfectly integrated, then we could find no spill-over even if such effects exist at the national level. In this sense, this study may a priori be seen as biased toward finding small or no general equilibrium effects of industrial composition even if they are present.

⁵ For example, even the seemingly large event of a city losing an industry that employed 10% of the workforce and paid a premium of 20% relative to other industries (roughly the situation facing Pittsburgh with the loss of the steel industry in the 1980s) implies only a 2% drop in the average wage using the pure accounting approach. Our result says that the total impact on city average wages would be a 6 to 7% decline. This is in line with Blanchflower and Oswald [1995]), who argue that the wage curve they identify fits with a number of models, including a bargaining-based model.

tralized bargaining. This result is important in its own right since it suggests that labor markets may involve considerable frictions even in an economy like the U.S., which is widely seen as highly competitive. While this class of models is often cited as helping understand unemployment, we believe that our results provide very strong evidence in support of this type of model for understanding wage behavior in the U.S. and potentially other market economies.

It is important to emphasize that we are interested in longer term differences in wage levels associated with different industrial composition as opposed to short run adjustments to industrial change and the related change in the level of employment. In particular, we focus mainly on changes in industrial composition that are orthogonal to changes in city level employment rates. Thus, our focus is different from, for example, that in Greenstone and Moretti [2003] or Blanchard and Katz [1992]. While Greenstone and Moretti emphasize shorter run demand effects by focussing mainly on local real estate price and same-industry wage bill effects within three years of acquiring a large plant, we focus on changes arising over 10 year horizons and are investigating whether there are wage effects of changes in industrial composition holding direct demand effects, or in other words the employment rate, constant. This focus also differentiates our work from studies of regional adjustment to demand changes such as in Blanchard and Katz [1992]. In order to clarify this difference, we take care in our empirical work to control for the types of demand effects examined in these papers.

Our empirical results are also related to a set of papers which examine the causes of city level employment and wage growth. In that literature, strong city performance has been variously linked to city size, the diversity of employment in a city [Glaeser, Kallal, Scheinkman, and Shleifer 1992], and the concentration of educated workers in a city [Moretti 2004]. In our empirical work, we introduce measures capturing each of these effects and show that these do not change our main findings. Thus, whatever we are identifying, it is over and above these other hypothesized driving forces. Our paper is also related to the voluminous literature aimed at understanding the effects of international trade on wages since much of this literature has debated the potential effects of trade-induced changes in industrial composition. The paper most closely related to ours is Borjas and Ramey [1995], which uses city level variation similar to ours to examine how trade induced changes in industrial composition may have affected returns to skill. Our focus is, nevertheless, very different since we focus on wage levels rather than on returns to skill.

The remaining sections of the paper are as follows. In Section 1, we present a model with

matching frictions and bargaining to illustrate how changes in industrial composition in an economy can affect wage setting within all sectors in the economy - including wages in sectors not directly involved in any composition shift. In Section 2, we use the model to derive a general empirical specification which embeds alternative views about the determination of wages. In particular, our empirical specification allows us to examine whether the data supports the view that general equilibrium effects from industrial composition changes are small and that the accounting approach is a justifiable procedure. In Section 3 we discuss the data used in the study and report basic empirical results. In Section 4 and 5 we address issues related to endogeneity, selection, robustness and interpretation. Section 6 concludes.

1 A Simple Model of the Spill-Over Effects of Industrial Composition.

In this section, we illustrate how a shift in the composition of jobs in a local economy can affect wages within all sectors in the economy, including sectors that are not part of the shift. We discuss this in the context of a general equilibrium search and bargaining model, showing that such a composition effect is separate from wage effects arising from shifts in the overall level of demand in the economy. For example, if a city experiences the loss of a high wage paying industry, one would expect an effect on local wages due to a reduction in overall demand. However, we show that, even holding the employment rate constant, a standard bargaining model also implies that the resulting shift in the composition of employment has an impact on wages in all sectors in the economy. Whether that composition effect is sizable is, of course, an empirical issue. The goal of this section is to derive an empirical strategy to explore whether changes in job composition have important general equilibrium (or spill-over) effects on within sector wage determination.

In order to illustrate the mechanisms for why job composition may affect the determination of wages, we consider an economy with one final good, denoted Y , which is an aggregation of output from I industries as given by

$$Y = \left(\sum_{i=1}^I a_i Z_i^\sigma \right)^{1/\sigma}. \quad (1)$$

The price of the final good is normalized to 1, while the price of the good produced by industry i is given by P_i . In this economy, we assume that there are C local markets called cities and that the industrial goods can be produced in any of these markets. The total quantity of the industrial good Z_i produced across the economy is equal to the sum across

cities of X_{ic} , the output in industry i in city c .

To simplify the exposition, we begin with an economy without worker mobility across cities. We extend the model in the next subsection to allow for worker mobility. To create a job in industry i in city c , a firm must pay a cost of c_{ic} , the value of which will be endogenously determined in equilibrium. If a job is filled, it generates a flow of profits for the firm given by

$$p_i - w_{ic} + \epsilon_{ic},$$

where w_{ic} is the wage and ϵ_{ic} is a city-industry specific cost advantage and $\sum_c \epsilon_{ic} = 0$. If we denote by V^f the discounted value of profits from a filled position, and we denote by V^v the discounted value of a vacancy, then V^f and V^v must satisfy the standard Bellman relationship given by

$$\rho V_{ic}^f = (p_i - w_{ic} + \epsilon_{ic}) + \delta(V_{ic}^v - V_{ic}^f), \quad (2)$$

where ρ is the discount rate and δ is the exogenous death rate of matches. If a firm does not fill a job it faces a per-period cost of r_i to maintain the position. Thus, the discounted value of profits from a vacant position must satisfy

$$\rho V_{ic}^v = -r_i + \phi_c(V_{ic}^f - V_{ic}^v), \quad (3)$$

where ϕ_c is the probability a firm fills a posted vacancy. For simplicity, and without loss of generality, we set $r_i = 0$.

Workers in the economy can be either employed or unemployed in a given period. The discounted value of being employed in industry i in city c , denoted U_{ic}^e , must satisfy the Bellman equation:

$$\rho U_{ic}^e = w_{ic} + \delta(U_{ic}^u - U_{ic}^e), \quad (4)$$

where U_{ic}^u represents the value associated with being unemployed when the worker's previous job was in industry i . When an individual is unemployed, he gets flow utility from an unemployment benefit, b , plus a city specific amenity term, τ_c .⁶ Letting ψ_c represent the probability that an unemployed individual finds a job, and $1 - \mu$ represent the probability the an individual finding a job gets a random draw from jobs in all industry versus being

⁶ We have not included an amenity term in the flow utility when employed to simplify notation. No results would change if we included such a term since what is important for wage determination is the difference in utility between being unemployed or employed. The amenity term τ should therefore be interpreted as the difference in flow utility associated with amenities when unemployed versus when someone is employed.

assigned a match in his previous industry,⁷ the value associated with being unemployed satisfies:

$$\rho U_{ic}^u = b + \tau_c + \psi_c(\mu U_{ic}^e + (1 - \mu) \sum_j \eta_{jc} U_{jc}^e - U_{ic}^u). \quad (5)$$

The important aspect to note from equation (5) is that, as long as $\mu < 1$, the utility level associated with having lost one's job in industry i depends on the utility associated with jobs in other industries. The instantaneous probability of finding a job in industry j is given by $\psi_c(1 - \mu)\eta_{jc}$, where η_{ic} represents the fraction of vacant jobs that are in industry j . This implies that a worker finding a job in another industry is assumed to find it in proportion to the relative size of that industry.

Once a match is made, workers and firms bargain a wage, which is set according to the Nash bargaining solution:

$$(V_{ic}^f - V_{ic}^v) = (U_{ic}^e - U_{ic}^u) \times \kappa, \quad (6)$$

where κ is a parameter governing the relative bargaining power of workers and firms. The probability a match is made is determined by the matching function:

$$M((L_c - E_c), (N_c - E_c)),$$

where L_c is the total number of workers in city c , E_c is the number of employed workers (or matches) in city c , and $N_c = \sum_i N_{ic}$ is the number of jobs in city c , with N_{ic} being the number of jobs in industry i in city c . Then, given the exogenous death rate of matches, δ , and assuming that the match function is of Cobb-Douglas form, the steady state condition is given by

$$\delta ER_c = M((1 - ER_c), (\frac{N_c}{L_c} - ER_c)) = (1 - ER_c)^\sigma (\frac{N_c}{L_c} - ER_c)^{1-\sigma}, \quad (7)$$

where ER_c is the employment rate. It follows that the proportion of jobs in industry i , η_{ic} , can be expressed as $\eta_{ic} = \frac{N_{ic}}{\sum_i N_{ic}}$.

⁷ We could easily add to an unemployed worker's trajectory the possibility of drawing an offer from another city. For example, if, with probability Γ , an unemployed worker from industry i in city c sampled jobs in industry i from other cities, then the value functions would need to satisfy:

$$\rho U_{ic}^u = b + \tau_c + (1 - \Gamma)\psi_c(\mu U_{ic}^e + (1 - \mu) \sum_j \eta_{jc} U_{jc}^e - U_{ic}^u) + \Gamma(\sum_c \psi_c \frac{N_{ic'}}{\sum_c N_{ic'}} U_{ic'}^e - U_{ic}^u),$$

where $N_{ic'}$ is the number of jobs in industry i in city c' . Since the additional term $\sum_c \psi_c \frac{N_{ic'}}{\sum_c N_{ic'}} U_{ic'}^e$ would act as a common element for workers across all cities, it would not change the implications we are focusing on which relate to within city effects. For this reason, we can disregard this possibility without loss.

The number of jobs created in industry i in city c , N_{ic} , is determined by the free entry condition:

$$c_{ic} = V_{ic}^v. \quad (8)$$

The cost c_{ic} should be viewed as the cost of creating a marginal job. If this cost were fixed, then cities would generally specialize in only one industry. Since we want to generate cities with employment across a wide range of industries, we assume that c_{ic} is increasing in the number of new jobs being created locally in that industry, which, in equilibrium, is proportional to N_{ic} . We also want to allow cities to have a comparative advantage in creating certain types of jobs relative to others. We therefore assume that c_{ic} is a decreasing function of the industry-city specific measure of advantage denoted Ω_{ic} . For tractability, we assume that the relationship between c_{ic} , N_{ic} and Ω_{ic} is given by:

$$c_{ic} = \frac{N_{ic}}{L_c(\Upsilon_i + \Omega_{ic})}.$$

In the above equation we have also assumed that c_{ic} is proportional to $\frac{N_{ic}}{L_c}$, in order to avoid scale effects. This assumption can easily be relaxed. The term Υ_i reflects any systematic differences in cost of entry across industries, which allows us to assume that $\sum_c \Omega_{ic} = 0$. Finally, the probability an unemployed worker finds a match and the probability a firm fills a vacancy satisfy

$$\psi_c = \frac{\delta ER_c}{1 - ER_c} \text{ and } \phi_c = \left(\frac{1 - ER_c}{\delta ER_c}\right)^{\frac{\sigma}{1-\sigma}}, \quad (9)$$

respectively.

At the city level, the price of industrial output is taken as given and an equilibrium consists of values of N_{ic} , w_{ic} , and ER_c that satisfy equations (6), (7) and (8). Note that these equilibrium values will depend upon (among other things) the city specific productivity parameters Ω_{ic} and ϵ_{ic} . An equilibrium for the entire economy has the additional requirement that the prices for industrial goods must ensure that markets for these goods clear.

At the city level, the equilibrium outcomes will reflect how a city adjusts its production mix to take advantage of the different prices of the industrial goods in relation to its comparative technological advantages in producing the goods. At the economy wide level, the change in prices will be caused by changes in demand for the industry level goods, captured in the a_i s. As we will make clear, our focus will be on isolating the effect of a change in job composition on city level wages. To this end, we will take the above description of a steady state equilibrium as representing an equilibrium at a point in time, and we will compare how this equilibrium changes in response to changes in the exogenous driving forces a_i , Ω_{ic} , and ϵ_{ic} .

1.1 The Interaction between Sectoral Wages and the associated Reflection Problem

Our focus will be on the determination of wages, as implied by (6). Our goal is to highlight how sectoral level wages inherit a strategic complementarity property in the presence of wage bargaining that gives rise to potentially rich general equilibrium effects of changes in industrial composition on wages. For now we will treat η_{ic} and ER_c as given, leaving the discussion of their endogenous determination and its implications for later.

To understand the forces determining wages in an industry-city cell, we begin by using equations (4) and (5) to express the value of finding a job relative to being unemployed as

$$U_{ic}^e - U_{ic}^u = \frac{w_{ic} - b - \tau_c}{\rho + \delta + \psi_c \mu} - \frac{\psi_c(1 - \mu) \sum_j \eta_{jc}(w_{jc} - b - \tau_c)}{(\rho + \delta + \psi_c \mu)(\rho + \delta + \psi_c)}. \quad (10)$$

A key feature of equation (10) is that, as long as $\mu < 1$, a worker's utility from being employed relative to being unemployed is a decreasing function of the wages paid in other sectors of the city's economy. This arises because higher wages elsewhere in the local economy imply a greater value of staying unemployed and finding a job in another sector. Also note that a compositional change captured by a change in the η s, holding ψ_c constant, will in general affect the utility of finding a job. For example, if the composition change implies more concentration of jobs in high wage sectors this decreases the value of a match in sector i .

To express the value of a match to a firm, we can use equations (2) and (3) in a similar fashion:

$$(V_{ic}^f - V_{ic}^v) = \frac{p_i - w_{ic} + \epsilon_{ic}}{\rho + \delta + \phi_c}. \quad (11)$$

From equations (10) and (11), we can use equation (6) to solve for w_{ic} . This is given by

$$w_{ic} = \gamma_{c0} + \gamma_{c1} p_i + \gamma_{c2} \sum_j \eta_{jc} w_{jc} + \gamma_{c1} \epsilon_{ic}, \quad (12)$$

where the coefficients in (12) are $\gamma_{c0} = \frac{(\rho + \delta + \phi_c) \kappa}{(\rho + \delta + \psi_c \mu) + \kappa(\rho + \delta + \phi_c)} (b + \tau_c)$, $\gamma_{c1} = \frac{\rho + \delta + \psi_c \mu}{(\rho + \delta + \phi_c) \kappa + (\rho + \delta + \psi_c \mu)}$ and $\gamma_{c2} = \frac{(\rho + \delta + \phi_c) \kappa}{[(\rho + \delta + \psi_c \mu) + \kappa(\rho + \delta + \phi_c)]} \frac{\psi_c(1 - \mu)}{(\rho + \delta + \psi_c)}$. Note that these coefficients are implicitly functions of the employment rate and that $0 < \gamma_{c2} < 1$ as long as $0 < \mu < 1$. Equation (12) captures the notion that in a search and matching framework sectoral wages act as strategic complements; that is, high wages in one sector are associated with high wages in other sectors. The strength of this strategic complementarity is captured by γ_{c2} . It is interesting to notice the effect of workers' bargain power on the size of γ_{c2} . As can be verified, γ_{c2} is an increasing in κ , implying that sectoral wages are more strongly positively linked the

weaker is workers' bargaining power. This suggests that even in environments where workers have minimal bargaining power, it is possible that γ_{c2} is high and that wages act as strong strategic complements. Notice also that equation (12) indicates why wages in one industry can vary across cities even when the productivity of the job is identical. In the case where $\mu = 1$ (i.e. workers are immobile across sectors), this effect disappears and wages are determined solely by value of marginal product. In such a case there would not be any effects of changes in industrial composition on within-sector wages. In this sense, the model nests more standard formulations where there are no spill-over effects. In contrast, with $\mu < 1$, a change in industrial composition can have an effect on within sector wages, even when the employment rate is unchanged.

If there is a pure industrial composition shift that causes a 1% increase in the average city wage, $\sum_j \eta_{jc} w_{jc}$, equation (12) indicates that within industry wages would then increase by $\gamma_{c2}\%$. But this is just a first round effect. Since the initial compositional change would affect all within industry wages, it would cause the average wage to increase by another γ_{c2} percent, inducing a further round of adjustments. Hence, there are feedback dynamics which continue to multiply themselves out. The total effect of the change in industrial composition on the average wage would therefore be $\frac{1}{1-\gamma_{c2}}$.⁸

Equation (12) has the structure of the classic reflection problem (Manski [1993], Moffitt [2001]) in that an outcome, here a sectoral wage, depends on the average of wages in a city. To make progress toward estimating such a relationship, it is necessary to overcome the simultaneity inherent to this type of interaction. To this end, we begin by rewriting equation (12) so that the sectoral wage is expressed only as a function of nationally determined prices and of the exogenous productivity terms ϵ :

$$w_{ic} = d_{ic} + \left(\frac{\gamma_{c2}}{1-\gamma_{c2}}\right)\gamma_{c1} \sum_j \eta_{jc}(P_j - P_1) + \gamma_{c1}\epsilon_{ic} + \gamma_{c1}\frac{\gamma_{c2}}{1-\gamma_{c2}} \sum_j \eta_{jc}\epsilon_{jc}, \quad (13)$$

where $d_{ic} = \gamma_{c0}\left(1 + \frac{\gamma_{c2}}{1-\gamma_{c2}}\right) + \gamma_{c1}\frac{\gamma_{c2}}{1-\gamma_{c2}}P_1 + \gamma_{c1}P_i$. In equation (13), we have expressed prices in relation to the price of an arbitrarily chosen good denoted P_1 in order to help emphasize how a pure shift change in industrial composition affects wages (where by a pure shift we mean a change in the η s that does not change the total number of jobs). From equation (13)

⁸ In order to get an upper bound on this effect, one can consider the limit case where the bargaining power of workers goes to zero (κ goes to infinity) and ρ goes to zero. In this case, $\gamma_{c2} = \frac{\theta_c(1-\mu)}{(\delta+\psi_c)}$. Using the definition of θ_c , this implies that $\frac{1}{1-\gamma_{c2}}(1-\mu) = \frac{1-\mu}{1-ER_c}$. If $1-ER$ is assumed to be captured by the average unemployment rate, say 4%, and workers stay in the same industry 80% of the time, then $\frac{1}{1-\gamma_{c2}}(1-\mu) = 5$. This indicates that a change in industrial composition that would have a direct effect of 1% on the average wage due simply to the composition shift, would have a total effect of 5% due to the strategic complementarity of wages. In this case, the spill-over effect would be four times the direct effect.

we see that a city with an industrial composition biased toward goods with high value added will have higher wages within each sector. This relationship indicates that industry wage differentials ultimately arise from differences in value added (which themselves arise because of differences in the demand parameters, a_i , in equation (1)), interacting with increasing costs of creating jobs. If we had not set the costs of a vacancy r_i to zero, these would also be a determinant of wage premia.

We can now use equation (12) or (13) to relate the price differential $P_i - P_1$ to the average national level wage premium in industry i relative to the a baseline industry 1. This is achieved by noting that $w_{ic} - w_{1c}$ is equal to $\gamma_{c1}(P_i - P_1) + \gamma_{c1}(\epsilon_{ic} - \epsilon_{1c})$. If we now take the average of $w_{ic} - w_{1c}$ across cities, we obtain

$$w_i - w_1 = \gamma_1(P_i - P_1) + \hat{d}_i, \quad (14)$$

where $w_i - w_1$ is the national level wage premium in industry i relative to to industry 1, γ_1 is the average of γ_{c1} across cities, and \hat{d}_i is an industry specific constant.⁹ Substituting (14) into (13), we get

$$w_{ic} = \tilde{d}_{ic} + \left(\frac{\gamma_{c2}}{1 - \gamma_{c2}}\right)\left(\frac{\gamma_{c1}}{\gamma_2}\right) \sum_j \eta_{jc}(w_j - w_1) + \gamma_{c1}\epsilon_{ic} + \gamma_{c1}\frac{\gamma_{c2}}{1 - \gamma_{c2}} \sum_j \eta_{jc}\epsilon_{jc}, \quad (15)$$

where $\tilde{d}_{ic} = d_{ic} + \left(\frac{\gamma_{c2}}{1 - \gamma_{c2}}\right)\left(\frac{\gamma_{c1}}{\gamma_1}\right) \sum_j \eta_{jc}\hat{d}_{ic}$. Equation (15) provides a direct expression of how wages within an industry-city cell depend on the industrial composition of a city as captured by the index $\sum_j \eta_{jc}(w_i - w_1)$. We will denote this index by R_c and refer to it as average city rent. A high value of this index indicates that a city's employment is concentrated in high paying industries. Thus, the specific composition effect captured in (15) is one related to the proportion of good jobs in a city.

In interpreting the effect of $R_c = \sum_j \eta_{jc}(w_i - w_1)$ on wages from equation (15), we are performing a partial equilibrium exercise as we treat the employment rate in a city (as reflected in the γ parameters), and the sectoral composition, as given. In order to capture the dependence of wages on the city's employment rate more explicitly, it is useful to take a linear approximation of (15) around the point where cities have identical industrial composition ($\eta_{ic} = \eta_i$) and employment rates ($ER_c = ER$), which arises when $\epsilon_{ic} = 0$ and $\Omega_{ic} = 0$. For simplicity, we actually perform the expansion at the point, where the η_i s are equal. Furthermore, to eliminate the city level fixed effects driven by the amenity, τ , it is useful to focus on the difference in wages within a city-industry cell across two steady state equilibria,

⁹The industry specific constant \hat{d}_i is equal to $\sum_c(\gamma_{c1} - \gamma_1)(\epsilon_{ic} - \epsilon_{1c})$. When we later take a first order approximation around an equilibrium where the ϵ s equal zero, this term will be equal to zero.

denoted Δw_{ci} . This is given by equation (16):

$$\Delta w_{ic} = \Delta d_i + \left(\frac{\gamma_2}{1-\gamma_2}\right)\Delta \sum_j \eta_{jc}(w_i - w_1) + \gamma_{i5}\Delta ER_c + \Delta U_{ic}, \quad (16)$$

where Δd_i is an industry specific effect ($\Delta d_i = \gamma_1 \frac{\gamma_2}{1-\gamma_2} \Delta P_1 + \gamma_1 \Delta P_i$) that does not vary across cities, and hence could be captured in an empirical specification by including industry dummies, and $\Delta U_{ic} = \gamma_1 \Delta \epsilon_{ic} + \gamma_1 \frac{\gamma_2}{1-\gamma_2} \sum_j \frac{1}{I} \Delta \epsilon_{jc}$ is the error term, with I being the total number of industries. In (16), the γ coefficients are the same as presented after equation (12), except now they are evaluated at common match probabilities, θ and ϕ . The added coefficient, γ_{i5} , reflects the effect of a change in the employment rate on wage determination and it depends on all the parameters of the model. In particular, even in a first order approximation, this coefficient may vary across industries since the effects of a tighter labor market may affect the bargaining power of firms in an industry with a high value added product differently from the bargaining power of firms in an industry with a low value added product.

The coefficient that interests us most in equation (16) is the coefficient on $\Delta R_c = \Delta \sum_j \eta_{jc}(w_j - w_1)$. If we could estimate this coefficient consistently, we would obtain an estimate of the extent of (structural) city-level strategic complementarity between wages in different sectors. This would be done by inferring the value of γ_2 from the estimated effect of changes in R_c . However, the coefficient $\frac{\gamma_2}{1-\gamma_2}$ is of interest in its own right as it provides an estimate of the total - direct and feedback - effect of a change of industrial composition on within sector wages, as opposed to γ_2 , which provides the partial uni-directional effect. Thus, if we examine wages in the same industry (say the textile industry) in different cities, a positive value for $\frac{\gamma_2}{1-\gamma_2}$ implies that those wages will be higher in cities where employment is more heavily weighted toward what we call high rent industries, where high rent industries are defined in terms of national level wage premia. This arises in the model because the workers in that industry have a better outside option to use when bargaining with firms in a higher average rent city. It is worth noting once again that this average rent effect is estimated conditioning on the employment rate. Thus, it is concerned with the composition of employment rather than the level of labor demand.

The coefficient $\frac{\gamma_2}{1-\gamma_2}$ relates directly to the question raised in the introduction concerning the extent to which the simple ‘‘composition adjustment’’ accounting procedure commonly used in the literature to calculate the effects of industrial composition on wages is appropriate. The coefficient $\frac{\gamma_2}{1-\gamma_2}$ provides an answer to this by quantifying the spill-over effects as multiples of the accounting procedure. To see this, note that the accounting procedure would calculate

the effect on the average wage within a city of a change in industrial composition as being given by the change in R_c .¹⁰ Given this, the total effect of a composition change is the accounting measure plus $\frac{\gamma_2}{1-\gamma_2}$ times that measure (or, $\frac{1}{1-\gamma_2}$ times that measure). If $\frac{\gamma_2}{1-\gamma_2}$ is estimated to be zero then the accounting procedure completely captures the effects of the composition shift. In terms of the model, this would fit with $\mu = 1$, a scenario in which there are no spill-overs. If, instead, $\frac{\gamma_2}{1-\gamma_2}$ is not zero, it indicates by what percentage the accounting measure needs to be multiplied by to capture the full effect of a change in industrial composition on the average wage.

Finally, as always, the potential to identify model parameters depends on the properties of the error term in the key estimating equation, which is our case equation (16). Given that this error term is $\Delta U_{ic} = \gamma_1 \Delta \epsilon_{ic} + \gamma_1 \frac{\gamma_2}{1-\gamma_2} \sum_j \frac{1}{I} \Delta \epsilon_{jc}$, this effectively reduces to properties of ϵ_{ic} . To discuss identification issues, it is helpful to express ϵ_{ic} (or Ω_{ic}) as the sum of a common component, which reflects absolute advantage, and a second component which captures relative advantage. For example, let $\hat{\epsilon}_{ct}$ represent the common component of the ϵ s and let v_{ict}^ϵ represent the relative advantage component, with $\epsilon_{ict} \equiv \hat{\epsilon}_{ct} + v_{ict}^\epsilon$, where the v_{ict} s

¹⁰ To see this more clearly, note that the standard accounting approach involves first estimating wage equation that controls for individual characteristics such as education, experience and industry. Then one recovers the estimates on industry dummies, denoted ν_{it} . These estimated industry coefficients correspond to the inter-industry wage differentials. Then, the accounting approach consists of computing the term

$$A_{ct} = \sum_i \nu_{it} [\eta_{ict+1} - \eta_{ict}],$$

which shows how the average wage in a city changes with the change in industrial composition when using a given set of industry wage premia. If a change in the local industrial composition of employment does not alter within industry wages then A_{ct} is a reasonable way to measure the impact of industrial change on the average wage. However, if wages are determined instead according to Equation (16), then we must include the impact of changes in local composition on local wages and the total effect of a change in industrial composition on the average wage in a city becomes:

$$\sum_i \nu_{it} [\eta_{ict+1} - \eta_{ict}] + \frac{\gamma_2}{1-\gamma_2} (R_{ct+1} - R_{ct}) = A_{ct} + \frac{\gamma_2}{1-\gamma_2} (R_{ct+1} - R_{ct}).$$

Interestingly, the change in R_{ct} is very closely related to A_{ct} given that the change in R_{ct} can itself be written as

$$R_{ct+1} - R_{ct} = A_{c,t} + \sum_i (\nu_{it+1} - \nu_{it}) \eta_{ict+1},$$

Hence the average wage change in a city can be expressed as

$$\left(1 + \frac{\gamma_2}{1-\gamma_2}\right) A_{c,t} + \frac{\gamma_2}{1-\gamma_2} \sum_i (\nu_{it+1} - \nu_{it}) \eta_{ict+1}.$$

Thus, in the case where ν is not varying over time, the change in $R_{c,t}$ is exactly the amount the accounting approach attributes to changes in industrial composition and the complete effect of a composition shift is one plus the total estimated spill-over effect times the simple accounting measure.

across industries within a city by definition sum to zero. As we shall see, identification of the parameters of equation (16) will depend primarily on properties of the absolute advantage component $\hat{\epsilon}_{ct}$. A rather minimal assumption that can be placed on the common component $\hat{\epsilon}_{ct}$ is that it behaves as a random walk; that is, that the increments in $\hat{\epsilon}_{ct}$ are independent of the past. Since this appears to us as a reasonable assumption, we will examine whether it is possible to consistently estimate equation (16) under this assumption recognizing that industrial structure and employment are endogenous. In particular, we will devise two instrumental variable strategies which will allow us to estimate (16) consistently under the sole assumption that the city-level common component of the ϵ s behave like a random walk. Since this assumption is nevertheless questionable, we will also devise a test of whether this condition holds. A more stringent, but still possible condition on the ϵ s is that the absolute advantage term be independent of the all relative advantage terms (present and past). This implies that whatever drives general city performance is not related to a particular pattern of industrial structure. While this assumption is clearly more demanding, we will show that it is sufficient for the OLS estimation of equation (16) to give consistent estimates. By comparing OLS and IV estimates we will therefore be able to evaluate whether the data are supportive or not of this more demanding assumption.

1.2 Endogeneity of Industrial Composition

Although our use of economy wide wage premia in (15) and (16) provides a means of resolving the reflection problem inherent to local interactions, this does not resolve all endogeneity issues. Let us begin by focusing on the potential endogeneity of ΔR_c , acting for now as if there were no endogeneity issue related to ΔER_c . The issue that arises in this case is that changes in R_c may be correlated with the error term in (16) due to the fact that industrial composition, as captured by η_{ic} , is correlated with the ϵ s. To see this potential endogeneity issue, we can use equation (7) to get the following expression for η_{ic} :

$$\eta_{i'c} = \frac{(\Upsilon_i + \Omega_{i'c})[p_i - d_{ci'} + (1 - \gamma_{c1})\epsilon_{i'c} + (\frac{\gamma_{c2}}{1-\gamma_{c2}})\gamma_{c1}(\sum_j \eta_{jc}(P_j - P_1) - \sum_j \eta_{jc}\epsilon_{jc})]}{\sum_i (\Upsilon_i + \Omega_{ic})[p_i - d_{ci} + (1 - \gamma_{c1})\epsilon_{ic} + (\frac{\gamma_{c2}}{1-\gamma_{c2}})\gamma_{c1}(\sum_j \eta_{jc}(P_j - P_1) - \sum_j \eta_{jc}\epsilon_{jc})]}. \quad (17)$$

We can express (17) more simply from taking a linear approximation:¹¹

$$\eta_{ic} \approx \frac{1}{I} + \pi_1(\epsilon_{ic} - \frac{1}{I} \sum_j \epsilon_{jc}) + \pi_2(P_i \Omega_{ic} - \frac{1}{I} \sum_j P_i \Omega_{ic}), \quad (18)$$

¹¹ In deriving (17), we are again taking the linear approximation around the point where $\epsilon_{ic} = 0$ and $\Omega_{ic} = 0$, which implies that the linear approximation is taken around a point where the employment shares and the employment rate are identical across cities. For simplicity, we again assume the common shares equal $\frac{1}{I}$.

where the π s represent positive coefficients obtained by the linear approximation.

The requirement for OLS to give consistent estimates of the coefficients in (16) can be expressed as follows:

$$\lim_{C,I \rightarrow \infty} \frac{1}{I} \frac{1}{C} \sum_{i=1}^I \sum_{c=1}^C \Delta R_c \Delta U_{ic} = \lim_{C,I \rightarrow \infty} \frac{1}{I} \frac{1}{C} \sum_{c=1}^C \Delta R_c \sum_{i=1}^I \Delta U_{ic} = 0, \quad (19)$$

where $\Delta R_c = \Delta \sum_j \eta_{jc}(w_i - w_1)$ and $\Delta U_{ic} = \gamma_1 \Delta \epsilon_{ic} + \gamma_1 \frac{\gamma_2}{1-\gamma_2} \sum_j \frac{1}{I} \Delta \epsilon_{jc}$, the error term in (16).¹² In Appendix A, we show that this condition is met under the assumption stated at the end of the previous section; that the absolute advantage component of the ϵ s in all periods are independent of the relative advantage components of the ϵ s and of the Ω s. This can be seen from Equation (19) since the city specific dimension of the η_{ic} s depends on relative advantage components, while $\sum_{i=1}^I \Delta U_{ic}$ depends only on the absolute advantage component of the ϵ s. If this condition is met, one need not worry about any endogeneity issue that may involve the wage premia ($w_i - w_1$) since the later do not vary with c . In intuitive terms, there must be no link between whether a city has experienced a shift in industrial composition toward high paying industries (a shift in ΔR) and other, general improvements affecting wages in all industries in the city. If this were not true, then the estimated effect of ΔR would partly reflect changes in the general conditions in a city rather than just capturing the impact of compositional shifts.

While we believe that the assumption ensuring consistency of the OLS estimates is possible, we are more comfortable with the much weaker assumption that the city-level common component of the ϵ has increments independent of the past. For this reason we will pursue two different instrumental variable strategies to address potential correlation between $\Delta \sum \eta_{ic}(w_i - w_1)$ and the error term in (16) that are valid under this weaker assumption. The idea behind one of our instrumental variables for $\Delta \sum \eta_{ic}(w_i - w_1)$ is to use national level observations on changes in industrial employment shares to predict city level industrial shares. This approach can be seen as using elements contained in the second term in (18), that is, in $\Delta(P_i \Omega_{ic} - \frac{1}{I} \sum_j P_j \Omega_{ic})$, to predict changes in η_{ic} . In Appendix A, we show that this instrument and our alternate instrument provide consistent estimates under the condition that $\Delta \hat{\epsilon}_c$ is independent of v_{ict}^ϵ and Ω_{ict} ; i.e., that changes in the absolute advantage of a city are independent of the past level of comparative advantage for the various industries in a city. This is a weaker assumption than that required for consistency under OLS because it allows for the possibility that a city that shifts in the direction of a higher wage industrial composition is also a city experiencing improvements in general - those generally improve-

¹² Where appropriate, the symbol \lim should be interpreted as a probability limit.

ments must be independent of past comparative advantage. While this assumption can also be placed in question, our use of two instrumental variable strategies will allow to test the identification strategy. We leave the details of this discussion and the construction of our instrumental variables for $\Delta \sum \eta_{ic}(w_i - w_1)$ until we get to the empirical section.

The last endogeneity issue relates to the potential correlation between the change in the employment rate, ΔER_c , and the error term in (16). To see this possibility, it is helpful to use a linear approximation of (7) and (8) to obtain an expression for ER_c in terms of the ϵ s and the Ω s. This is given in Equation (20), where the coefficients $\tilde{\pi}$ are again positive and obtained from the linear approximation:

$$ER_c \approx \tilde{\pi}_0 + \tilde{\pi}_1 \sum_j \epsilon_{jc} + \tilde{\pi}_2 \sum_j P_j \Omega_{ic}. \quad (20)$$

From (20) one can see that ΔER_c is likely correlated with the error term in (16) and, as a result, that OLS will generate inconsistent estimates. We will address this with an instrumental variable strategy in which we again use national level information on growth patterns to predict city level changes in employment rates. The approach is similar to that used by Blanchard and Katz [1992] in a closely related problem. Details of the construction of this instrument are again left to the empirical section.

1.3 Worker Mobility

In the model as presented so far, we have assumed that workers are not mobile across cities. At first blush, it may appear that the result that wages can vary systematically across cities due to the composition of employment will disappear once we allow for mobility of workers. However this will not necessarily be the case even if we allow for some directed search across cities.¹³ To see this, consider adding births and deaths to the previous model and have workers choose, after birth, a city in which to live so as to maximize their expected utility, taking in to account housing prices and local amenities. In particular, assume that workers care about wages, the price of housing in a city, p_{ct}^h , and about a local amenity, Ψ_c . In this case, a worker's (indirect) flow utility when employed in industry i in city c could be expressed as, $w_{ic} - \varsigma p_{ct}^h + \Psi_c$. Accordingly, his or her flow utility when unemployed will be given by $b + \nu_c - \varsigma p_{ct}^h + \Psi_c$. Note that in such a setup, wage negotiation will not be directly affected by housing prices since it is a cost that is incurred whether or not someone is employed. However, housing prices will need to adjust to equilibrate expected utility across cities. To capture this role, housing prices can be expressed as a positive function of

¹³ Allowing for random search across cities will not significantly change our previous analysis.

the population of a city and of amenities, such as

$$p_{ct}^h = d_0 + d_1 L_{ct} + d_2 \Psi_c.$$

It is straightforward to show that in this world, a city with a higher employment rate and a better employment mix, as captured by a higher value of R_{ct} , will attract more workers. This immigration will drive up local housing prices, causing the migration to stop before wages are equalized across cities. In fact, housing prices will adjust such that a city with a favorable composition of the jobs (due to favorable ϵ s and Ω s) induces local benefits that are captured by local landowners. Nonetheless, the impact of a favorable job composition will be seen in wages, as implied by equations (12)-(16), since these equations are still valid given that mobility decisions take place prior to wage negotiations. Although such worker mobility does not change the prediction of the model in terms of the complementarity of wages across sectors and the implied spill-over effects of industrial composition, it does have implications for house prices and labor flows across cities. In particular, it implies that the determinants of wages emphasized in equation (16) should also be determinants of housing prices and inter-city migration flows. We will therefore investigate, toward the end of the paper, whether changes in R_{ct} have these predicted impact on inter-city migration flows and on housing prices.

1.4 Worker heterogeneity

Up to now, labor in our model has been treated as homogenous and not differentiated by skill. Since workers in reality do differ in skills (and are treated differently by employers as a result), we need to take this into account when exploring the relevance of the spill-over effects predicted by the model. In particular, our model should be seen as applying to labor markets defined by skills. Empirically, we can approach the issue of insuring we are only estimating within skill-group implications of the model in two ways. The first is to work with sub-samples of the data intended to be homogeneous with respect to skill (e.g., using only young high school graduates). The other is to eliminate skill differentials by first controlling for a flexible set of “skill” indicators in the determination of wages (where a wage regression would now be seen as an individual level wage regression rather than one specified at the industry \times city level as it has been to this point). This later approach is justified if the setup costs and benefits associated with employing a more skilled worker are proportional to the increased productivity of the worker and if all the parameters of the model are identical across skill groups. To implement this approach, we would also need to control for the same

set of skill indicators when calculating the national level wage premia used in constructing the R_{ct} measure. We pursue both approaches in the empirical section.

2 Empirical Implementation

Our baseline empirical specification is given by Equation (21), which is a simple rewrite of equation (16) where we have divided both sides by w_1 in order to focus on a log specification, and where we have added a time subscript since we will pull data from different periods.

$$\Delta \log w_{ict} = \alpha_{1i} d_{it} + \alpha_2 \Delta R_{ct} + \alpha_{3i} \Delta ER_{ct} + \Delta U_{ict} \quad (21)$$

In (21) the d_{it} s are time varying industry dummies, $\alpha_2 = \frac{\gamma_2}{1-\gamma_2}$ is our main coefficient of interest, $R_{ct} = \sum_j \eta_{jc} (\frac{w_i}{w_1} - 1)$ is our index of industrial composition, α_{3i} are the coefficients capturing the effect of city level employment rates on wages (where these effects are allowed to vary across industries), and ΔU_{ict} is the error term defined by $\Delta U_{ict} = \frac{\gamma_1}{w_1} \Delta \epsilon_{ict} + \frac{\gamma_1}{w_1} \frac{\gamma_2}{1-\gamma_2} \sum_j \frac{1}{I} \Delta \epsilon_{jct}$. Our goal is to investigate the null hypothesis that $\alpha_2 = 0$ or, in other words, whether disregarding inter-sectoral wage spill-over provides an appropriate description of wage determination in local economies. Support for this null hypothesis would indicate that the standard accounting procedure completely captures the effect of a local change in the composition of employment on the local average wage. Our alternative hypothesis is that $\alpha_2 > 0$. A finding of $\alpha_2 > 0$ would indicate the presence of a spill-over from industrial composition to wages as predicted by the model and would indicate that the standard accounting approach is an inappropriate means of evaluating the effects of changes in industrial composition on average wages.

When estimating the effect of R_{ct} on wages, it is appropriate to worry about omitted variable bias, especially given existing alternative explanations for differences in wages across cities such as those related to city size, education levels (Moretti [2004], Acemoglu and Angrist [1999]), and diversity of employment in a city [Glaeser, Kallal, Scheinkman, and Shleifer 1992]. To allow for such issues, we will add to equation (21) measures related to these explanations as additional covariates, $Z_{c,t}$. As we suggested previously, we also address the potential issues of endogeneity of $R_{c,t}$ and $ER_{c,t}$. In addition, we will address the potential for worker mobility to cause a selection bias.

3 Data and Basic Results

3.1 Data

The data we use in the following investigations come from the 1970, 1980, 1990 and 2000 U.S. Census Public Use Micro-Samples (PUMS). We focus on wage and salary earners, aged 20 to 65 with positive weekly wages who were living in a metropolitan area at the time of the Census. To form our dependent variable we use the log of weekly wages, calculated by dividing wage and salary income by annual weeks worked (we also verified the robustness of our results to using hourly wages). We create real wages (in 1990 dollars) using the national level CPI as the deflator. Given our use of multiple Censuses, an important part of our data construction is the creation of consistent definitions of cities, education groups and industries over time. We provide the details on how we address these issues in Appendix B.

As we described in the previous section, one approach to addressing worker heterogeneity is to control for observable skills in a regression context. Our actual approach is to use a common two-stage procedure. In the first stage, we run individual level regressions of log wages using all the individuals in our national sample on categorical education variables (4 categories), a quadratic in experience, interactions of the experience and education variables, a gender dummy, black, hispanic and immigrant dummy variables, and the complete set of interactions of the gender, race and immigrant dummies with all the education and experience variables. We run these regressions separately by Census year to allow for changes in returns to skills over time. The regressions also include a full set of industry-by-city cell dummies and it is the coefficients on those that are used to construct the dependent variable in the second stage regression (equation (21) above). We eliminate all industry-city cells with fewer than 20 included individuals in any of the years. We use the square root of the number of observations in each industry-city cell to form weights for the second stage estimation. For most of our estimates, we use decadal differences within industry-city cells for each pair of decades in our data (1980- 1970, 1990-1980, 2000-1990), pooling these together into one large dataset and including period specific industry dummies. In all the estimation results we calculate standard errors allowing for clustering by city and year.

The main covariate in our estimation is the ΔR_{ct} variable which is a function of the national industrial wage premia and the proportion of workers in each industry in a city. We estimate the wage premia in a regression at the national level in which we control for the same set of education, experience, gender, race and immigration variables described for our first stage wage regression and also include a full set of industry dummy variables. This regression is

estimated separately for each Census year. The coefficients on the industry dummy variables are what we use as the wage premia in constructing our R measures.

3.2 OLS Results

We begin our presentation of results with the estimates of (21) without the inclusion of any additional control (Z_{ct}) variables. The first two columns of Table (1) contains the results from OLS estimation of the regression. These two regressions and all of those that follow include a full set of time varying industry dummy variables (3×144), thus allowing for changes in industry premia over time, but we do not present the long list of corresponding coefficients here. In column (1) we restrict the coefficient on the employment rate to be equal across industries. In column (2) we allow for the employment rate to vary across large industry groups.¹⁴ The coefficient on the change in R_{ct} variable is 2.6 in both specifications and is statistically significantly different from zero at any conventional significance level. If OLS provides consistent estimates of this coefficient, the fact that this coefficient is both economically substantial and statistically significant implies a rejection of the null hypothesis that the impact of changes in the composition of employment in a city is completely captured in the standard accounting measure. Further, the coefficient fits with the alternative hypothesis that cities with employment structures that shift toward higher premia industries have better wage performance within industries. Recall from our discussion of the definition of the R_{ct} variable that the magnitude of the coefficient on this variable can be interpreted as a multiple of the standard accounting effect. Thus, the OLS estimate implies that the total effect on average wages of a shift in composition toward higher paying industries is approximately three and a half times what is reflected in a standard accounting measure. This total effect may initially sound overly large but it is worth recalling that the accounting measure effects tend to be quite small. For example, let us consider the average real weekly wage for men with a BA or higher education (examples with other education or gender groups give similar results). For this group, the average wage increased by 8% across the cities in our sample between 1980 and 1990. If we recalculate the 1990 average wage for this group holding the industrial composition constant, the increase becomes 7%, implying that the accounting measure of the impact of shifts in industrial composition is 1%. Our estimates suggest that the total impact of shifts in industrial composition would

¹⁴ We are allowing the effect of the employment rate to vary across 16 industry groups. We also explored the effect of allowing employment to vary with all 144 industries, and found similar results. Since in most of our specifications, we could not reject that 16 interaction were sufficient, we adopted this specifications as our baseline case.

be 3.5% in this example. Such an increase is certainly larger than what is usually attributed to industrial shifts but is still only just over 40% of the overall increase. The fact that direct accounting measures of the impact of industrial shifts tend to be small has led to a discounting of explanations for changes in the U.S. wage structure that might show up through such shifts. Trade, for example, is usually relegated to a lower place in the list of potential explanations for this reason. An estimate of the size we report may imply that there is reason to re-examine those types of explanations.

One point of interest about this result is whether it is being driven by a subset of cities, such as those that faced particularly large re-adjustment after the difficulties in key manufacturing industries. To examine this, in Figure (1) we plot the change in city average wages (the average of our dependent variable across industries within a city) against ΔR_{ct} .¹⁵ The key point from this figure is that there is a strong positive relationship between wage changes and changes in our R measure that is not driven by outliers.

We are also interested in whether the estimated effect stems from some particular set of industries for which wages are particularly sensitive to the presence of high premia industries. We investigate this by re-estimating our basic specification interacting the ΔR variable with a complete set of industry dummy variables. This is equivalent to re-writing the α_2 coefficient with an i subscript. In Figure (2), we present a histogram of the full set of these α_{2i} coefficients. What is noteworthy in this figure is the concentration of values around the mean. The implication is that workers in virtually all industries benefit from a shift in employment composition toward high paying sectors and benefit to much the same degree. It is worth recalling when considering this result that we estimate industry premia while controlling for observable skills. Thus, high premium industries are not necessarily high skill industries.

It is interesting to note that our OLS estimates of α_2 in Table 1 are very stable whether or not we allow the employment rate effects to vary across industries (a pattern that also holds true when we move to estimating the relationship by instrumental variables). In fact, the OLS estimate of α_2 is almost identical if we don't include the change in the employment rate in the regression. The reason for this can be easily understood from Figure 3, in which we provide a plot of the change in the employment rate against ΔR_{ct} . The flat relationship in this figure indicates that these two variables are almost orthogonal to each other. This implies that most of the shifts in industrial composition look like pure compositional shifts, with little co-movement in a city's employment rate.

¹⁵We actually first regress $\Delta \log w_{ict}$ on industry-year dummies and plot the weighted average of the residuals from that regression in order to obtain a plot that replicates our actual regression.

4 Addressing Endogeneity and Selection Issues

4.1 Endogeneity: Methods and Results

As we discussed earlier, OLS estimation of (21) will provide consistent estimates if changes in a city’s absolute advantage is independent of patterns (levels and changes) of relative advantage. While this may be the case, in this section we want to explore the estimations (21) using two instrumental variable strategies which rely on weaker identification assumptions. To help motivate our instrumental variable strategies, it is useful to start by writing out a standard decomposition of ΔR_{ct} :

$$\Delta R_{ct} = \sum_i \Delta \eta_{ict} \nu_{it} + \sum_i \eta_{ict} \Delta \nu_{it}, \quad (22)$$

where $\nu_{it} = \frac{(w_{it}-w_{1t})}{w_{1t}}$ the national level industrial premium. Thus, movements in R_{ct} are decomposed into a component due to changes in the composition of the workforce (in the η_{ict} s) holding the wage premia constant and a component due to changes in the premia holding the local composition of the workforce constant.

Our instrumenting strategy for ΔR consists of building two instruments that emphasize variation in different components of the above decomposition, and which by construction are not functions of the $\Delta \epsilon$ s. The first is constructed using the following procedure. We first predict a level of employment for industry i in city c in period $t + 1$ using the formula:

$$\hat{l}_{ict+1} = l_{ict} \left(\frac{l_{it+1}}{l_{it}} \right).$$

That is, we predict future employment in industry i in city c using the employment in that industry in period t multiplied by the growth rate for the industry at the national level. Using these predicted values, we construct a set of predicted industry specific employment shares, $\hat{\eta}_{ict} = \frac{\hat{l}_{ict}}{\sum_i \hat{l}_{ict}}$, for the city in period $t + 1$ and form a measure given by:

$$IV1_{ct} = \sum_i \nu_{it} (\hat{\eta}_{ict+1} - \eta_{ict}) \quad (23)$$

This variable is closely related to the first term in the decomposition of the R measure given in (22). Thus, this instrument isolates the variation in ΔR that stems from changes in the employment composition, but instead of using actual employment share changes we use predicted changes based on national level changes, breaking the direct link between city level employment and wage changes. Essentially, $IV1$ focuses attention on the question, “what is the impact on local wages of a national level demand shift (stemming from, for

example, trade or preference shocks) if that shift is distributed across cities according to start of period employment shares?” Recall that use of this type of variation is implied by the model, where shifts in national level demand ($\{a_i\}$) are translated into local shifts in employment shares because of local differences in comparative advantage that will be reflected in initial period employment shares. In terms of the econometrics, since $IV1$ is not a function of $\Delta\epsilon_{ict}$, it is not correlated with the error term in (21), under the less demanding identification assumption that the change in the common component of the ϵ_{ict} (the absolute advantage component) is not predictable based on past information on comparative advantage. This contrasts with the condition for consistency of OLS which requires that changes in the absolute advantage component are also not correlated with changes in comparative advantage. A test of equivalence of the IV and OLS estimates is, therefore, a test of the stronger assumptions required for consistency of the OLS estimator. Our second instrument is designed to isolate the variation inherent in the second term in the decomposition, (22): the variation stemming from changes in wage premia over time, weighted by the importance of the relevant industry in the local economy. Thus, our second instrument is given by:

$$IV2_{ct} = \sum_i \eta_{ict}(\Delta\nu_{it}).$$

This instrument may initially seem less natural, since the discussion to this point has been almost entirely couched in terms of shifts in the concentration of employment. However, if our theoretical explanation, which emphasizes bargaining power, is correct then it should not matter whether the average premium available in the city declines because a high paying industry shuts down or because the premium paid in that industry declines. In either case, workers in other industries end up with a less valuable outside option. This would imply that we should get similar results using $IV1$ and $IV2$. Hence, examining whether the results obtained using these two alternative instruments are the same can provide a means of evaluating whether the outside option effect outlined in Section 2 is the likely mechanism at play. As with $IV1$, $IV2$ is uncorrelated with the error term in (21) under our main identifying assumption that the common component in the $\Delta\epsilon$ s is independent of the past comparative advantage.

Both instruments perform well in the first stage estimation. The F-statistic from the test of the significance of $IV1$ in the first stage regression of ΔR on the instrument takes a value of 15.8 and has an associated p-value of 0.0. The same statistic for $IV2$ is 31.6 with a p-value of 0.

A non-zero covariance between ΔER_c and ΔU_{ic} is to be expected since a city with a set of

large increases in its local, industry specific cost shocks (its ϵ_{ic} s) will have higher employment. We respond to that endogeneity problem using an instrument that is similar to IV1.¹⁶ In particular, we use as an instrument $\sum_i \eta_{ic} g_i$, where g_i is the growth rate of employment in industry i at the national level. Thus, the instrument is the weighted average of national level industrial employment growth rates, where the weights are the start of period industrial employment shares in the local economy. A city that has a strong weight on an industry that turns out to grow well at the national level will have a high value for this instrument. Because the ϵ_{ic} s are local demand shocks that sum to zero across cities, their movements are not correlated with the g_i s by construction. Finally, under the assumption that $\Delta\epsilon_{ic}$ is independent of ϵ_{ic} , the changes in ϵ_{ic} that constitute ΔU_{ic} will be independent of the ϵ_{ic} s used as weights in the instrument, resulting in a zero correlation between the instrument and the error term.

We present results from instrumental variables estimation using *IV1* and *IV2* individually, in the third and fourth columns of Table (1). In all the instrumental variable estimates we present, the employment rate is treated as endogenous and instrumented using $\sum_i \eta_{ic} g_i$. The third column contains results from instrumental variable estimation in which we use *IV1* to counter the potential endogeneity of ΔR_{ct} . The estimated coefficient of ΔR_{ct} is very similar to that obtained from OLS estimation and is again highly statistically significant. The fourth column contains results when we use *IV2*, the instrument that uses changes in industry premia over time. The similarity in the estimated coefficients obtained using *IV1* and *IV2* is striking; an outcome which we have argued fits well with theories of the impact of R that are based on changes in bargaining power. The results in columns 3 and 4 of Table (1) indicate that both moves away from high paying jobs and reductions in the premia associated with high paying jobs have approximately the same impact on within industry wages as implied by our bargaining story. In Column 5 we use both *IV1* and *IV2* simultaneously as instruments.

The identification assumption underlying our instrumental variable estimation of Equation (21), although plausible, can certainly be placed into question. Interestingly, both our instrumental variables *IV1* and *IV2* rely on the same assumption since the instruments are essentially different weighted averages of a city's initial employment shares and our key identifying assumption relates to those shares. Given that the weighting schemes are different, one would expect that if the identification assumption does not hold, these two IV strategies would produce different estimates since they would weight any departures from the assump-

¹⁶ This instrument is similar to that used in Blanchard and Katz [1992].

tion differently. This intuition can actually be formalized by considering a test of whether the estimates of α_2 using *IV1* versus *IV2* give similar results. The idea of the test is similar to a Hausman test: under the null both are consistent and should provide similar estimates. The only issue is that the variance-covariance matrix for the differences in estimates based on the two instruments needs to be calculated differently from a standard Hausman test since neither IV estimator is efficient. We perform this test,¹⁷ and not surprisingly given the similarity in the estimates in Column 3 and 4 in Table (1) or (2), it passes easily. We believe that the fact that the two IV approaches, which focus on very different data variation, give very similar results provides considerable support for the assumption that the common city level component in the ϵ acts like a random walk. Moreover, given that OLS also gives similar results suggests, further, that even the strong assumption needed for the consistency of OLS likely holds in these data.¹⁸

4.2 Selection: Methods and Results

A second key concern in estimating (21) is with selection of workers across cities. The R variable varies at the city level over time. Thus, changes in unobserved skills in a city that are correlated with movements of R will imply a non-zero coefficient on R that does not reflect general equilibrium effects of the type we are considering. For example, suppose that there are unobserved skills (which we will call ability) and that high premia industries can choose higher ability workers from lines of applicants. Suppose, further that the most able workers move out of a city if it loses a high paying industry, regardless of the industry in which they themselves are employed, because they want to live in a place where they have a chance of getting into a higher paying job. In that case, shifts in R may actually pick up the effects of shifts in the unobserved ability distribution.

We address selection concerns in a number of ways. First, we control for observable skill variables (education and experience) both when estimating the wage premia in the national level wage regression and when obtaining the industry-city average wages that form our key dependent variable. Our second approach is to implement the selection correction estimator that Dahl [2002] proposes and implements in his examination of regional variation in returns to education.

To understand the nature of Dahl’s approach, consider a model in which each worker has

¹⁷ We also performed a more standard over-identification test for column 5 and have not found any evidence against our maintained assumption.

¹⁸ The conditions that allow OLS to provide consistent estimates of (21) can also be shown to imply that estimating the relationship in levels instead of in differences should give similar results. Interestingly, when we estimate (21) in level we actually get similar results.

a (latent) wage value that he would earn if he lived in each possible city and chooses to live in the city in which his wage net of moving costs is highest. If we explicitly introduce individual heterogeneity, this implies that we should write the regression corresponding to observed wages as

$$E(\log w_{kict} | d_{kct} = 1) = \alpha_{0t} + \beta_{1t} x_{kct} + \alpha_1 ER_{ct} + \alpha_2 R_{ct} + \nu_i + \nu_c + E(e_{kct} | d_{kct} = 1), \quad (24)$$

where k indexes individuals and d_{kct} is a dummy variable equaling one if worker k is observed in city c at time t . The last, error mean, term is non-zero if worker city selection is not independent of the unobserved component of wages. If one were to estimate equation (21) not taking account of this error mean term then the estimated regression coefficients will suffer from well-known consistency problems.

In situations such as the union wage premium literature where there are only two options facing a worker, it is well known that the error mean term can be expressed as a function of the probability of selecting the given option [Heckman 1979, Lee 1983]. In our case, with multiple possible destinations to choose from, the error mean term will potentially be a function of characteristics of all of them, making estimation complicated. Dahl [2002] argues that under specific sufficiency conditions, the error mean term is only a function of the probability that a person born in the same state as k would make the choice that k actually made, greatly simplifying the problem. In his examination of the impact of selection of location across states on returns to education, however, he argues that this sufficiency assumption is overly restrictive and that one can effectively account for selection using functions of the probability k did not move from his state of birth and the probability he moved to the state in which he is observed at the time of the Census. Following work such as Ahn and Powell [1993] and Heckman and Robb [1985] for the binary choice case, he also proposes a non-parametric estimator for the relevant probabilities and the function of them that enters the regression of interest. We follow his approach with a few adjustments to account for the facts that we include immigrants in our analysis and that we are dealing with cities. Details on our selection estimation are provided in Appendix C. In essence, this estimator identifies the error mean (selection) effect using differences in the probabilities of being observed in a given city between two people who are identical in education, experience, race and gender but are born in different states. The idea is that, for example, people born in Oregon are more likely to be observed in Seattle than people born in Pennsylvania because Oregon is so much closer. If both are in fact observed living in Seattle then we are assuming that the person from Pennsylvania must have a larger Seattle specific “ability” (a stronger earnings related reason for being there) and this is what is being captured when we include functions of the

relevant probabilities of being observed in Seattle for each of them. Identification in this approach is based on the exclusion of state of birth by current city of residence interactions from the wage regression. That is, we assume that being born in a state close to your city of residence (or, more generally, a state with a high associated probability of moving to that city) does not directly determine the wage a worker receives.¹⁹

In practical terms, this approach to the potential selection problem again involves two estimation steps. In the first, as before, we estimate individual level regressions of log wages on the same complete set of education and experience variables, indicators for race, immigrant status, and gender, as well as a full set of city-by-industry dummies but now also add our proxies for the error mean term. We again retain the coefficients on the city-industry dummy variables and then proceed with the second stage regressions as before. The coefficients on the error mean proxy variables are jointly highly significant in the first stage regressions, implying that there are significant sample selection issues being addressed with this estimator.

In Table (2), we recreate the results from Table (1) while implementing Dahl (2002)'s selection correction. The resulting estimates for α_2 are very similar in magnitude to those obtained when we did not correct for sample selection. Thus, we do not believe that movements in unobserved ability across cities is strongly contaminating our estimates. The implication of the selection analysis is that while workers do select themselves across cities in a manner that is non-random with respect to earnings outcomes, changes in their selection pattern are not correlated with changes in the average industrial premium paid in a city.

4.3 Observing Strategic Complementarity

The results from the previous section suggest that changes in industrial composition toward higher paying jobs have the effect of increasing wages across all sectors. The key mechanism that leads to the amplification of an initial industrial composition change on wages in our model is that wages across sectors play the role of strategic complements. This effect was expressed most clearly in Equation (12). As previously emphasized, Equation (12) implies that a change in industrial composition that initially increases average wages in a city by 1%, would lead to a cumulative increase in average wages by a factor of $(\frac{1}{1-\gamma_2})\%$ due to the strategic complementarity of wages. Given that α_2 in Equation (21) relates to γ_2 according to $\alpha_2 = \frac{\gamma_2}{1-\gamma_2}$, our estimates of α_2 that are around 2.5 in Table (1) suggest that γ_2 should

¹⁹Note that this is different from assuming that state of birth does not affect current wages since, even if we include a set of state of birth dummy variables in our first stage estimation, our approach remains identified off interactions between city-of-employment and state-of-birth.

be in the range of $0.71 = (\frac{2.5}{3.5})$. This implication of the model can also be examined directly by estimating Equation (12) by instrumental variables. Since the coefficients in Equation (12) depend on the employment rate, we proceed as with Equation (15) to take a linear approximation of (12) around the point where the ϵ s and Ω s are zero and then taking first differences, to get a linear equation of the form

$$\Delta \ln w_{ict} = \psi_1 d_{it} + \psi_2 \Delta \sum_j \eta_{jct} \ln w_{jct} + \psi_{3i} \Delta ER_{ct} + \tilde{U}_{ict} \quad (25)$$

where ψ_2 corresponds to the γ_2 in the model, and the error term corresponds to $\gamma_1 \Delta \epsilon_{ict}$. In the case of Equation (25), estimation by OLS would definitely be expected to give upward biased estimates of $\psi_2 = \gamma_2$ since the relationship suffers from the reflection problem. However, it can be verified that instrumental variable estimation of Equation (25) using our previous set of instruments should give consistent estimates under the same assumption as before, that is, under the assumption that the common component of the ϵ s (a city's absolute advantage) is independent of the past. It is worth emphasizing that the difference between Equation (21) and (25) pertains only to the main variable of interest. In (25) this variable is the average city wage,²⁰ while in (21) it is a city level average of national wage premia.

Estimates of Equation (25) that parallel Tables (1) and (2), are present in Tables (3) and (4). In all cases we control for industry specific effects of the employment rate and a full set of time varying industry dummies as before; the difference between Tables (3) and (4) being that in Table (4) we implement the selection correction. The first thing to note from these two tables is that, as one should expect, there is now a large and significant difference between estimates of ψ_2 (denoted Δ Av. City Wage in the tables) obtained by OLS or IV. The OLS estimate is .853, which if translated to compare with α_2 would imply an $\alpha_2 = 5.80 = \frac{.853}{1-.853}$. However, in this case there are no conditions for which we should expect OLS to give consistent estimates. In contrast, when we estimate by IV, we get an estimate of ψ_2 equal approximately to .69, which implies a value for $\alpha_2 = 2.22 = \frac{.69}{1-.69}$, which is very close to that obtained Tables (1) and (2) using different approaches. In particular, the estimation of (21) by OLS provides one means of overcoming the reflection problem by focusing on national level wage premia, while the IV estimation of (25) provides a conceptually quite different approach. In the remain sections of the paper we focus on verifying the robustness of the results presented in Table (1) and (2). In order to save space, we do not provide further results on the estimation of (25), but we have verified that they are as robust as those for Equation (21). Also, in all

²⁰ Calculated as the city average of the regression adjusted city-industry wages that form our dependent variable.

subsequent sections of the paper, we present results incorporating Dahl’s sample selection correction.

5 Further Explorations of the Wage Premia Effects

5.1 Other Driving Forces for City Level Wage Changes

Ours is certainly not the first attempt to examine the determinants of city level wage changes and/or city-level growth. The literature on what makes for a high performing city has produced a number of hypotheses. In this section, we introduce measures corresponding to some of the more prominent hypotheses to see whether our R_{ct} measure may be spuriously capturing one of these alternative driving forces.

One prominent explanation for city level growth is provided in Glaeser *et al.* [1992]. They examine city level growth over time in the U.S., comparing the impact of measures of city size, which would be important determinants of growth if agglomeration type models were driving growth patterns, and measures of the industrial diversity of the economy. They argue that the importance of diversity is implied by, for example, Jane Jacob’s theorizing. They find that industrial diversity is a stronger determinant of city specific growth than city size. In response to this view, in the first three columns of Table (5), we introduce a measure of the “fractionalization” of employment in a city at the start of each decade. The measure of fractionalization we use is one minus the Herfindahl index, or one minus the sum of squared industry shares. This measure itself tends not to be very significant in our estimates and, more importantly, does not change our estimates of the α_2 coefficient.

Another possibility relates to the recent literature on education externalities which examines the claim that having a larger proportion of workers in a city being highly educated benefits all workers in the city. Moretti [2004], for example, in an examination of wages in U.S. cities in the 1980s finds that cities with a greater increase in the proportion of workers with a BA or higher education have higher wage gains. Acemoglu and Angrist [1999] find weaker results for the impact of education using average years of education in a state. Again, we are interested in whether our R_{ct} measure is actually picking up this alternative effect. It is worth re-emphasizing, though, that we control for education in the regressions from which we estimate our national level wage premia and, thus, the R_{ct} measure does not reflect cities that have high wages because they have high levels of education. In the middle set of three columns in Table (5), we introduce the change in the proportion of workers with a BA or higher education (the College Share) as an additional regressor. The college share

variable itself enters significantly, supporting Moretti [2004]’s findings, but introducing this variable has very little impact on our estimates of the effect of changes in R_{ct} .²¹ Although not reported, we also examined the the effect of using average years of education as an alternative measure of the education level of a city. This latter variable does not enter significantly, supporting results in Acemoglu and Angrist [1999] and fitting with the often contradictory results in this literature. Moreover, including average years of education does not affect the estimates of our R_{ct} effects.

Finally, in columns (7), (8) and (9), we introduce the change in the log of the size of the city’s labor force. This city size variable is intended to capture the type of agglomeration effects tested in, for example, Glaeser *et al.* [1992]. It can also be viewed as a direct control for overall demand effects, allowing us to check whether the effects of shifts in industrial composition we are measuring are truly pure compositional effects or whether they are partly capturing shifts in overall demand that may accompany the loss or gain of major industries. In the IV columns, we instrument for the labor force variable using the same type of instrument we have used in various ways to this point; that is, one based on predicting labor force growth in a city from national level growth for each industry combined with the initial industrial composition in the city. Whether instrumented or not, this variable has small and statistically insignificant effects and, more importantly, its inclusion does not alter the estimates of the ΔR effect. Again, this fits with our interpretation of α_2 as capturing a pure compositional effect.

Overall, our conclusion is that while some of the other hypothesized factors we have considered may affect city level wage growth, we are not inadvertently picking any of them up with our R_{ct} measure. Moreover, the impact of the shift in industrial composition toward higher paying industries is much stronger than any of the effects from these competing explanations.

5.2 Robustness Checks

One possible explanation for the patterns we observe is that while standard trade forces are affecting wages in tradeable goods sectors, wages associated with skills that are used in the non-tradeable sector are moving for standard demand-induced reasons. Thus, a shift in employment in a city toward having more workers with high levels of unobserved skills (perhaps because of their pursuit of local amenities) could lead to an increase in R_{ct} that would not affect wages in the tradeable sector for standard factor price equalization reasons

²¹It is worth noting, though, that Sand [2006] finds that this positive and significant impact is observed in the 1980s but not in the 1970s or 1990s when estimation is carried out separately by decade.

but, because the higher skilled workers have more income to spend on locally produced non-traded goods, it could affect wages in the non-traded sector. Under this explanation, we should see smaller impacts of changes in R_{ct} on wages in tradeable sector industries than on wages in non-tradeable sector industries.

To define tradeable and non-tradeable sectors, we rely on an approach suggested in Jensen and Kletzer [2005]. They argue that the share of output or employment in tradeable goods should vary widely across regional entities (cities in our case) since different cities will be more heavily concentrated in producing different goods which they can then trade. For non-tradeable goods, on the other hand, assuming that preferences are the same across cities, one should observe similar proportions of workers in their production across cities. We rank industries by the variance of their employment shares across cities and call the industries in the top third, high trade industries, those in the middle third, medium trade industries, and those in the bottom third, low trade industries.²² In Table (6), we present estimates of our basic model carried out separately for the low, medium and high trade industries. While the estimated effect of changes in R_{ct} do tend to be slightly higher for the low trade industries, the effects for the medium and high trade industries continue to be strongly significant and of the same order of magnitude as the estimated effects we obtained from the overall sample. Thus, our results do not appear to be arising simply because of spill-overs into the non-traded goods sector labor market. In Table (7) we also present results of estimating α_2 separately for twelve industry groupings. As can be seen, the results are very similar across these industries, with the effect estimated for manufacturing industries corresponding closely to that observed for the overall sample. This is further indication that the effects do not seem concentrated in non-trade good sectors since most manufactured goods are tradeable across cities.²³

5.3 Education Breakdowns

To this point, we have established that shifts in industrial composition toward higher paying sectors is associated with growth in wages in almost all industries, it holds up to corrections for endogeneity and sample selection, and it is not proxying for explanations for city growth based on overall demand, diversity of the industrial structure, or changes in the education level of the workforce. We are interested, now, in examining whether the effects are ubiqui-

²²The actual observations in the low trade industries is much lower than those in the medium and high trade industries because the low trade industries tend to be small and so tend to be disproportionately dropped when we impose our restriction that a given industry-city cell must contain at least 20 observations.

²³Interestingly, the sector where wage exhibit the least response to change in R_{ct} is public administration. This may be due to the fact that wage for federal employees are generally set at the national level.

tous across different education groupings. As previously indicated, the model presented in section 2 conceptually applies to workers of given skill group. Accordingly, in this section we report results associated with estimating equation (21) for different skill groups defined by education and experience. In particular, we consider three education groups: workers with at most a high school education, workers with some post secondary education but without a BA, and finally a group with at least a BA. We further divide each of these groups into young workers (those with less than 10 years of experience²⁴) and older workers (those with more than ten years of experience). For each of these groups we calculated an R_{ct} variable that is specific to the group.²⁵ The estimates of the effects of changes in R_{ct} on each of these groups is presented in Table (8). All the results in this table also control for changes in the employment rate interacted with 16 industry grouping as before.²⁶

The results in Table (8) indicate substantial, though not identical, effects of shifts in industrial composition on within-industry wages in all the skill groups. The estimates reveal a slight education gradient, with ΔR effects being smaller for those with a BA or more in both experience groups. However, these differences are not very significant. The main dimension on which there are significant differences is with respect to experience. As can be seen, the effects of a change in industrial composition on wages is much greater for younger workers than for older workers.²⁷ While this observation is not consistent with a strict interpretation of the model, it is potentially consistent with an internal labor market view in which more experienced workers tend to be partially sheltered from direct market comparisons. For example, if workers in ongoing relationship only renegotiate their wages infrequently, this is the pattern we would expect. Hence, we believe that the results based on focusing on different skill groups paint a similar picture as that documented previously, that is, changes in industrial composition appear to have substantial spillover effects throughout the labor

²⁴The measure of experience we use is potential experience, defined by Age - Years of School - 6. Appendix B contains details.

²⁵ We create the national level wage premia, and thus the R_{ct} measures, separately for each skill group. We, again, estimate in two steps, with the first step individual wage regression (including Dahl's selection correction) as well as the second step run separately for the 6 skill groups.

²⁶ For the results of this table, the effects of employment rates on different industries are constrained to be identical across the 6 skill groups. We also estimated the model without imposing such constraints. This gave similar, but less precise estimates.

²⁷ Since we use different R_{ct} measures for each skill group means that the group-specific results in Table (8) do not bear an easy mechanical relationship to the overall results in Table (2). This explains why the estimates for some groups in Table (2) that are substantially below the Table (2) estimates are not counter-balanced by substantially higher estimates for other groups. In contrast, if we use the single, overall R_{ct} measure for all groups, we find the same education and experience patterns as in Table (8) but with coefficients that are all greater. For example, the estimates for young workers with high school education is above 4, while the result for older post-secondary workers is above 2.

market.

5.4 Additional Effects Associated with Changes in Industrial Composition

Up to this point, our empirical investigation has focused on evaluating how shifts in industrial composition toward high paying jobs affects wages. In this subsection, we briefly explore the effects of such a change on other city level outcomes. Given our interpretation of the wage effects, it would be natural to expect - as documented in section 1.3 - that a change in industrial composition that favors high paying jobs should be associated with in-migration and, potentially, an increase in the price of housing. In Table (9) we investigate this possibility. In the first three columns of the table, we examine whether changes in R_{ct} are associated with increases in the price of housing, as measured by the rent for one bedroom apartments. As can be seen from the table, we again observe a positive association. It is worth noting that while the estimated coefficient on the change in R_{ct} varies substantially across our different estimation strategies, in all three cases we find that housing prices capitalize a large fraction of the changes in wages.²⁸

It is interesting, in addition, to consider the effect of changes in R_{ct} on labor force growth, as we do in the specifications in columns 4, 5 and 6 in Table (9). The results in these columns show that a change in industrial composition in favor of high paying jobs has a robust positive association with labor force growth. Together, the observations on the effects of shifts in industrial composition on labor force growth and housing costs suggest that a city that experiences a positive increase in R_{ct} becomes a more attractive city, as we would expect given the impact in terms of higher average wages that we demonstrated in the rest of the paper.

6 Conclusion

Policy forums often involve discussions of the effects and desirability of attracting or retaining jobs in high paying industries. In the popular press, it is common to hear statements claiming that economic success is closely tied to favoring employment growth in sectors that pay high wages for comparable individuals. In contrast, the most prevalent view among economic researchers is that changes in industrial composition generally contribute very little to labor

²⁸ Since rents makes up only a fraction of total consumption, under perfect mobility of workers across cities we could expect that the effect of a change in R_{ct} on housing prices would be greater than the effect on wages.

market performance and therefore a focus on the effects of different policies with respect to the creation or destruction of better paying jobs is likely misplaced. This consensus position is based primarily on evaluating the economic impact of changes in industrial composition using a simple accounting approach which assumes away spill-over effects from the loss of jobs in one sector on wages in other sectors. Although traditional economic theory may provide good reason to believe that such spill-over effects should be absent or small, in this paper we build on random matching models to highlight an empirical strategy for evaluating the spill-over effects of changes in job composition on industry level wage payments. We implement that strategy using U.S. census data from 1970 to 2000. Our main finding is that spill-overs appear pervasive, persistent and large. In particular, at the city level we find that having jobs more concentrated in high paying industries has an effect on the average wage within the city that is 2.5 to 4 times larger than that implied by the common composition adjustment accounting approach. We show that these results are robust to using different instrumental variable strategies, controlling for worker selection and focusing on sectors producing highly tradeable goods.

Our results suggest that policies or events which affect industrial composition should not be evaluated simply using the standard accounting approach but instead should explicitly take account of substantial general equilibrium effects. For example, it is common for the opening up of trade relationships to involve a reallocation of high and low paying jobs across trading partners. Our results suggest that a proper evaluation of the effects of increased trade needs to incorporate the potential spill-over effects on wages in other sectors. In general, recognizing and quantifying these feedback effects will lead to much more variable assessments of the gains from trade since markets that attract high paying industries will benefit more than traditionally thought, while markets that lose such jobs should benefit less.²⁹

²⁹ Beaudry, Green and Collard [2005] and Beaudry and Collard [2006] find that increased openness to international trade over the period 1978-98 had very uneven effects across countries. In particular, countries that attracted high-capital-high-wage industries gained dis-proportionally relative to countries that increased employment in low-capital intensive industries. The spill-over effects found in this paper offer a potential explanation for the size of the effects found in these two papers.

A Examining Consistency

As described in the text, we are interested in the condition:³⁰

$$\lim_{C, I \rightarrow \infty} \frac{1}{I} \frac{1}{C} \sum_{i=1}^I \sum_{c=1}^C \Delta R_c \Delta U_{ic}, \quad (\text{A1})$$

which, using $R = \sum_j \eta_{jc}(w_j - w_1)$, can be written as

$$\lim_{C, I \rightarrow \infty} \frac{1}{I} \frac{1}{C} \sum_{i=1}^I \sum_{c=1}^C \left[\sum_j \Delta \eta_{jc}(w_j - w_1) + \sum_j \eta_{jc} \Delta(w_j - w_1) \right] \Delta U_{ic}$$

or

$$\lim_{C, I \rightarrow \infty} \frac{1}{I} \frac{1}{C} \left[\sum_j (w_j - w_1) \sum_c \Delta \eta_{jc} \sum_i \Delta U_{ic} + \sum_j \Delta(w_j - w_1) \sum_c \eta_{jc} \sum_i \Delta U_{ic} \right]. \quad (\text{A2})$$

We will handle the limiting arguments sequentially, allowing $C \rightarrow \infty$ first. Then, we are concerned with 2 components in (A2), which we will handle in turn. The first is

$$\lim_{C \rightarrow \infty} \frac{1}{C} \sum_c \Delta \eta_{jc} \sum_i \Delta U_{ic}. \quad (\text{A3})$$

Given the decomposition $\epsilon_{ic} = \hat{\epsilon}_c + v_{ic}^\epsilon$, where $\sum_i v_{ic}^\epsilon = 0$, we get

$$\Delta \eta_{jc} = \pi_1(\Delta v_{jc}^\epsilon) + \pi_2(\Delta P_j \Omega_{jc} - \Delta P \bar{\Omega}_c), \quad (\text{A4})$$

where \bar{x}_c equals the simple average of x_{ic} across i within a city.

Also,

$$\sum_i \Delta U_{ic} = (\gamma_1 + \frac{\gamma_1 \gamma_2}{1 - \gamma_2}) I \Delta \hat{\epsilon}_c \quad (\text{A5})$$

Then, given that $E(\Delta \hat{\epsilon}_c) = 0$ (again, recalling that we have removed economy-wide trends) and if $\Delta \hat{\epsilon}_c$ is independent of Δv_{ic}^ϵ and $(\Delta P_j \Omega_{jc} - \Delta P \bar{\Omega}_c)$, it is straightforward to show that (A3) equals zero.

The second component is,

$$\lim_{C \rightarrow \infty} \frac{1}{C} \sum_c \eta_{jc} \sum_i \Delta U_{ic} \quad (\text{A6})$$

where $\sum_i \Delta U_{ic}$ is again given by (A5), while η_{jc} is given by (20). For (A6) to be zero we require in addition that $\Delta \hat{\epsilon}_c$ be independent of past values of v_{ic}^ϵ and of $(P_j \Omega_{jc} - P \bar{\Omega}_c)$. Thus,

³⁰Throughout this appendix we omit the t subscript for simplicity.

if $\Delta\hat{\epsilon}_c$ is independent of the past and is independent of Δv_{ic}^ϵ and $(\Delta P_j \Omega_{jc} - \Delta P \bar{\Omega}_c)$, then (A1) equals zero and OLS is consistent.

We are also interested in the conditions under which our instruments can provide consistent estimates. Apart from the instruments being correlated with ΔR_{ic} , the condition we require for a given instrument, Z_c is,

$$\lim_{C, I \rightarrow \infty} \frac{1}{I} \frac{1}{C} \sum_{i=1}^I \sum_{c=1}^C Z_c \Delta U_{ic} \quad (\text{A7})$$

For what we call IV1,

$$Z_c = \sum_j \eta_{jc} (g * _j - 1) (w_j - w_1),$$

where $g * _j = \frac{1+g_j}{\sum_k \eta_{kc} 1+g_k}$ and g_j is the growth rate in employment in industry j at the national level. Given this, (A7) becomes:

$$\lim_{C \rightarrow \infty} \frac{1}{C} \sum_j (w_j - w_1) \sum_c \eta_{jc} (g * _j - 1) \sum_i \Delta U_{ic}. \quad (\text{A8})$$

Thus, (A8) equals zero under the same conditions under which (A6) equaled zero, i.e., that $E(\Delta\hat{\epsilon}_c) = 0$ and $\Delta\hat{\epsilon}_c$ to be independent of past values of v_{ic}^ϵ and $(P_j \Omega_{jc} - P \bar{\Omega}_c)$. Obviously, this condition will be satisfied is $\hat{\epsilon}_c$ behaves as a random walk with increments independent of the past.

Similarly, the relevant condition when using IV2 is given by,

$$\lim_{C \rightarrow \infty} \frac{1}{C} \sum_j \Delta(w_j - w_1) \sum_c \eta_{jc} \sum_i \Delta U_{ic} \quad (\text{A9})$$

and the same conditions ($\Delta\hat{\epsilon}_c$ to be independent of past values of v_{ic}^ϵ and of $(P_j \Omega_{jc} - P \bar{\Omega}_c)$) ensure that this condition equals zero.

Several points follow from this discussion. First, OLS can provide consistent estimates and for it to do so requires the assumptions needed for the IV's to provide consistent estimates (that changes in the absolute advantage for a city are independent of the initial set of comparative advantage factors for that city) plus the stronger assumption that changes in absolute advantage and changes in comparative advantage are independent. Thus, if OLS and the two IV estimates are equal then this is a test of the stronger assumption about independence in changes. Second, if the key identifying assumption underlying the IV's is not true (i.e., changes in absolute advantage are not independent of past comparative advantage) then the two IV's weight the problematic correlation (between $\Delta\hat{\epsilon}_c$ and v_{ic}^ϵ) differently (in particular, IV1 weights using the weights $(w_j - w_1)$, while IV2 uses the weights $\Delta(w_j - w_1)$) and estimates based on the different IV's should differ.

B Data Construction

The Census data was obtained with extractions done using the IPUMS system (see Ruggles *et al.* [2004]). The files were the 1980 5% State (A Sample), 1990 State, and the 2000 5% Census PUMS. For 1970, Forms 1 and 2 were used for the Metro sample. The initial extraction includes all individuals aged 20 - 65 not living in group quarters. All calculations are made using the sample weights provided. For the 1970 data, we adjust the weights for the fact that we combine two samples. We focus on the log of weekly wages, calculated by dividing wage and salary income by annual weeks worked. We impute incomes for top coded values by multiplying the top code value in each year by 1.5. Since top codes vary by State in 1990 and 2000, we impose common top-code values of 140,000 in 1990 and 175,000 in 2000.

A consistent measure of education is not available for these Census years. We use indicators based on the IPUMS recoded variable EDUCREC that computes comparable categories from the 1980 Census data on years of school completed and later Census years that report categorical schooling only. To calculate potential experience (age minus years of education minus six), we assign group mean years of education from Table 5 in Park [1994] to the categorical education values reported in the 1990 and 2000 Censuses.

Census definitions of metropolitan areas are not comparable over time since, in general, the geographic areas covered by them increase over time and their definitions are updated to reflect this expansion. The definition of cities we use attempts to maximize geographic comparability over time and roughly correspond to 1990 definitions of MSAs provided by the U.S. Office of Management and Budget.³¹ to create geographically consistent MSAs, we follow a procedure based largely on Deaton and Lubotsky [2001] which uses the geographical equivalency files for each year to assign individuals to MSAs or PMSAs based on FIPs state and PUMA codes (in the case of 1990 and 2000) and county group codes (for 1970 and 1980). Each MSA label we use is essentially defined by the PUMAs it spans in 1990. Once we have this information, the equivalency files dictate what counties to include in each city for the other years. Since the 1970 county group definitions are much coarser than those in later years, the number of consistent cities we can create is dictated by the 1970 data. This process results in our having 152 MSAs that are consistent across all our sample years. Code for this exercise was generously provided by Ethan G. Lewis. Our definitions differ slightly from those in Deaton and Lubotsky [2001] in order to improve the 1970-1980-1990-2000 match. We use an industry coding that is consistent across Censuses and is based on the IPUMS

³¹See <http://www.census.gov/population/estimates/pastmetro.html> for details.

recoded variable IND1950, which recodes census industry codes to the 1950 definitions. This generates 144 consistent industries.³² We have also replicated our results using data only for the period 1980 to 2000, where we can use 1980 industry definitions to generate a larger number of consistent industry categories.³³ We are also able to define more (231) consistent cities for that period.

C Implementing the Selection Estimator

As described in the paper, our main approach to addressing the issue of selection on unobservables of workers across cities follows Dahl [2002]. Dahl argues that the error mean term in equation (24) for person j can be expressed as a function of the full set of probabilities that a person born in j 's state of birth would choose to live in each possible city in the Census year. Further, he presents a sufficiency assumption under which the error mean term is a function only of the probability of the choice actually made by j . That sufficiency condition essentially says that two people with the same probability of choosing to live in a given city have the same error mean term in their regression: knowing the differences in their probabilities of choosing other options is not relevant for the size of the selection effect in the process determining the wage where they actually live. Dahl, in fact, presents evidence that this assumption is overly restrictive and settles on a specification in which the error mean term is written as a function of the probability of making the migration choice actually observed and the probability that the person stayed in their birth state.

Implementing Dahl's selection correction approach requires two further decisions: how to estimate the relevant migration probabilities and what function of those probabilities to use as the error mean term. For the first, Dahl proposes a non-parametric estimator in which he divides individuals up into cells defined by discrete categories for education, age, gender, race and family status. He then uses the proportion of people within the cell that is relevant for person j who actually made the move from j 's birth state to his destination and the proportion who stayed in his birth state as the estimates of the two relevant probabilities. This is a flexible estimator which does not impose any assumptions about the distribution of the errors in the processes determining the migration choice. For the second decision, Dahl

³²See <http://usa.ipums.org/usa-action/variableDescription.do?mnemonic=IND1950> for details.

³³ The program used to convert 1990 codes to 1980 comparable codes is available at <http://www.trinity.edu/bhirsch/unionstats>. That site is maintained by Barry Hirsch, Trinity University and David Macpherson, Florida State University. Code to convert 2000 industry codes into 1990 codes was provided by Chris Wheeler and can be found at <http://research.stlouisfed.org/publications/review/past/2006>. See also a complete table of 2000-1990 industry crosswalks at <http://www.census.gov/hhes/www/ioindex/indcswk2k.pdf>

uses a series estimator to provide a non- parametric estimate of the error mean term as a function of these probabilities.

We essentially implement Dahl’s approach in the same manner apart from several small changes. First, we are examining the set of people who live in cities in the various Census years but we only know the state, not the city of birth. We form probabilities of choosing each city for people from each state of birth. People who live in a city in their state of birth are classified as “stayers” and those observed in a city not in their state of birth are classified as “movers”.³⁴ We estimate the error mean term as a function of the probability that a person born in j ’s state of birth moved to j ’s city of residence and the probability that a person born in j ’s state of birth still resided in that same state. Stayers have an error mean term which is a function only of the probability that the person stayed in their state of birth (since the probability of their actual choice and the probability of staying are one and the same).

As in Dahl [2002], we estimate the relevant probabilities using the proportion of people within cells defined by observable characteristics who made the same move or who stayed in their birth state. Similar to Dahl [2002], we define the cells using 4 education categories, 8 age categories, gender and a black race dummy. For stayers, we also use extra dimensions based on family status.³⁵ This is possible because of the larger number of stayers than movers. The full interaction of these various characteristics defines 80 possible person types for the movers and 240 for stayers. For the movers in a particular city (i.e., for the set of people born outside the city in which that city is situated), the probabilities will also differ based on where the person was born. Thus, identification of the error mean term comes from the assumption that where a person was born does not affect the determination of their wage, apart from through the error mean term. Intuitively, a person born in Pennsylvania has a lower probability of being observed in Seattle than a person born in Oregon. If both are in fact observed living in Seattle then we are assuming that the person from Pennsylvania must have a larger Seattle specific “ability” (a stronger earnings related reason for being there) and this is what is being captured when we include functions of the relevant functions of being observed in Seattle for each of them. For stayers, we do not have this form of variation and, hence, identification arises from the restriction that family status affects the decision to stay in one’s state of birth but not (directly) the wage.

Our main difference relative to Dahl [2002] is that while he drops immigrants, we keep them

³⁴For cities that span more than one state, we call a person who is observed in a city that is at least partly in their birth state a stayer.

³⁵Specifically, we use single, married without children, and married with at least one child under age 5.

in our sample. We essentially treat them as if they are born in a different state from the city of residence except that we do not include a probability of their remaining in their place of birth. We divide the rest of the world into 11 regions (or “states” of birth). As with other movers, we divide them into cells based on the same education, age, gender and race variables and assign them a probability of choosing their city of residence. Contrary to other movers, however, we do not assign them the probability that immigrants from their region of birth are observed in their own city in the current Census year. Instead, we assign them the probability that a person with their same education was observed in their city in the previous Census. This follows the type of ethnic enclave assumption used in several recent papers on immigration, essentially using variation based on the observation that immigrants from a particular region tend to migrate to cities where there are already communities of people with their background.

Having obtained the estimated probabilities of following observed migration paths and of staying in state of birth, we need to introduce flexible functions of them into our regressions. In practice, we introduce these functions in our first estimation stage. The specific functions we use are quadratics in the estimated probabilities. For movers born in the U.S., we introduce a quadratic in the probability of moving to the actual city from the state of birth and a quadratic in the probability of remaining in the state of birth. For stayers, we introduce a quadratic in the probability of remaining in the state in general. For immigrants, we introduce a quadratic in the probability that people from the same region and with the same education chose the observed city. This represents a restriction on Dahl [2002], who allowed for separate functions for each destination state. We, instead, assume the parameters in the functions representing the error mean term are the same across all cities.

References

- ACEMOGLU, D., AND J. ANGRIST (1999): “How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws,” NBER Working Papers 7444, National Bureau of Economic Research, Inc.
- AHN, H., AND J. L. POWELL (1993): “Semiparametric estimation of censored selection models with a nonparametric selection mechanism,” *Journal of Econometrics*, 58(1-2), 3–29.
- BEAUDRY, P., AND F. COLLARD (2006): “Globalization, Returns to Accumulation, and the World Distribution of Income,” *Journal of Monetary Economics*.

- BEAUDRY, P., F. COLLARD, AND D. GREEN (2005): “Demographics and Recent Productivity Performance: Insights from Cross-Country Comparisons,” *Canadian Journal of Economics*, (2), 309–344.
- BLANCHARD, O. J., AND L. F. KATZ (1992): “Regional Evolutions,” *Brookings Papers and Economic Activity*, (1), 1–75.
- BORJAS, G. J., AND V. A. RAMEY (1995): “Foreign Competition, Market Power, and Wage Inequality: Theory and Practice,” *Quarterly Journal of Economics*, pp. 1075–110.
- BOUND, J., AND G. JOHNSON (1992): “Changes in the Structure of Wages in the 1980’s: An Evaluation of Alternative Explanations,” *American Economic Review*, pp. 371–392.
- DAHL, G. B. (2002): “Mobility and the Return to Education: Testing a Roy Model with Multiple Markets,” *Econometrica*, (6), 2367–2420.
- DEATON, A., AND D. LUBOTSKY (2001): “Mortality, Inequality and Race in American Cities and States,” NBER Working Papers 8370, National Bureau of Economic Research, Inc.
- GLAESER, E., H. KALLAL, J. SCHEINKMAN, AND A. SHLEIFER (1992): “Growth in Cities,” *Journal of Political Economy*, (6), 1126–1152.
- GREENSTONE, M., AND E. MORETTI (2003): “Bidding for Industrial Plants: Does Winning a ‘Million Dollar Plant’ Increase Welfare?,” NBER Working Papers 9844, National Bureau of Economic Research, Inc.
- HECKMAN, J. J. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 47(1), 153–162.
- HECKMAN, J. J., AND R. J. ROBB (1985): “Alternative methods for evaluating the impact of interventions,” in *Longitudinal Analysis of Labor Market Data*, ed. by J. Heckman, and B. Singer. Cambridge University Press.
- JENSEN, B., AND L. KLETZER (2005): “Tradable Services: Understanding the Scope and Impact of Services Outsourcing,” Working Paper 05-9, Institute for International Economics.
- LEE, L. (1983): “Generalized Econometric Models with Selectivity,” *Econometrica*, 51(2), 507–512.

- MANSKI, C. F. (1993): “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies*, 60(3), 531–42.
- MOFFITT, R. A. (2001): “Policy Interventions, Low Level Equilibria, and Social Interactions,” in *Social Dynamics*, ed. by S. N. Durlauf, and H. P. Young. MIT Press.
- MORETTI, E. (2004): “Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data,” *Journal of Econometrics*, 121(1-2), 175–212.
- MORTENSEN, D. T., AND C. A. PISSARIDES (1999): “New developments in models of search in the labor market,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 39, pp. 2567–2627. Elsevier.
- OSWALD, A. J., AND D. G. BLANCHFLOWER (1995): *The Wage Curve*, vol. 1 of *MIT Press Books*. The MIT Press.
- PARK, J. H. (1994): “Estimation of Sheepskin Effects and Returns to Schooling Using the Old and the New CPS Measures of Educational Attainment,” Papers 338, Princeton, Department of Economics - Industrial Relations Sections.
- RUGGLES, S., M. SOBEK, T. ALEXANDER, C. A. FITCH, R. GOEKEN, P. K. HALL, M. KING, AND C. RONNANDER (2004): *Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]*. Minneapolis, MN: Minnesota Population Center [producer and distributor].
- SAND, B. M. (2006): “The Social Returns to Education,” University of British Columbia Working Paper.

Table 1: Basic Results

	OLS	OLS	IV1	IV2	IV1&IV2
Variables	(1)	(2)	(3)	(4)	(5)
ΔR_{ct}	2.580 (0.193)*	2.574 (0.193)*	2.940 (0.424)*	2.790 (0.4)*	2.846 (0.391)*
ΔER_{ct}	0.444 (0.089)*				
$\Delta ER_c \times \text{Ind.}$	No	Yes	Yes	Yes	Yes
Obs.	28425	28425	28425	28425	28425
R^2	0.55	0.552			

Notes: Each column is an estimate of equation (21). Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies.

Table 2: Basic Results with Selection Correction

	OLS	OLS	IV1	IV2	IV1&IV2
Variables	(1)	(2)	(3)	(4)	(5)
ΔR_{ct}	2.626 (0.201)*	2.619 (0.201)*	2.996 (0.451)*	2.879 (0.437)*	2.919 (0.425)*
ΔER_{ct}	0.433 (0.097)*				
$\Delta ER_c \times \text{Ind.}$	No	Yes	Yes	Yes	Yes
Obs.	27945	27945	27945	27945	27945
R^2	0.595	0.597			

Notes: Each column is an estimate of equation (21) using the selection correction procedure described in section (4.2) of the text. Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies.

Table 3: Reflection Specification

	OLS	IV1	IV2	IV1&IV2
Variables	(1)	(2)	(3)	(4)
Δ Av. City Wage	0.853 (0.013)*	0.689 (0.038)*	0.695 (0.038)*	0.692 (0.036)*
$\Delta ER_c \times$ Ind.	Yes	Yes	Yes	Yes
Obs.	28425	28425	28425	28425
R^2	0.626			

Notes: Each column is an estimate of equation (25). Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies. Average city wage is equal to $\sum_i \eta_{jct} \ln w_{ict}$, as described in equation (25).

Table 4: Reflection Specification with Selection Correction

	OLS	IV1	IV2	IV1&IV2
Variables	(1)	(2)	(3)	(4)
Δ Av. City Wage	0.857 (0.012)*	0.689 (0.039)*	0.698 (0.039)*	0.694 (0.038)*
$\Delta ER_c \times$ Ind.	Yes	Yes	Yes	Yes
Obs.	27945	27945	27945	27945
R^2	0.667			

Notes: Each column is an estimate of equation (25) using the selection corrected wages. Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies. Average city wage is equal to $\sum_i \eta_{jct} \ln w_{ict}$, as described in equation (25).

Table 5: Alternative Explanations

Variables	OLS	IV1	IV2	OLS	IV1	IV2	OLS	IV1	IV2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔR_{ct}	2.579 (0.192)*	2.875 (0.41)*	2.810 (0.409)*	2.534 (0.184)*	2.637 (0.391)*	2.573 (0.367)*	2.504 (0.209)*	2.862 (0.605)*	2.731 (0.572)*
1 - Herfindahl	0.274 (0.198)	0.276 (0.208)	0.274 (0.203)						
Δ BA or >				0.548 (0.157)*	0.563 (0.159)*	0.562 (0.16)*			
Δ log labour force							0.018 (0.015)	0.01 (0.031)	0.008 (0.031)
$\Delta ER_c \times$ Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	28425	28425	28425	28425	28425	28425	28425	28425	28425
R^2	0.553			0.557			0.553		

Notes: Each column is an estimate of equation (21) with an added regressor which reflects an alternative hypothesis regarding wage outcomes as discussed in section (5.1). Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies.

Table 6: By Trade and Non-Trade Industries

Variables	Low Trade			Medium Trade			High Trade		
	OLS	IV1	IV2	OLS	IV1	IV2	OLS	IV1	IV2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔR_{ct}	3.118 (0.399)*	3.960 (1.141)*	3.802 (0.846)*	2.723 (0.227)*	3.240 (0.638)*	2.934 (0.831)*	2.385 (0.169)*	2.702 (0.368)*	2.640 (0.341)*
$\Delta ER_c \times$ Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	4606	4606	4606	11754	11754	11754	12065	12065	12065
R^2	0.491			0.525			0.607		

Notes: Each column is an estimate of equation (21) estimated separately for low, medium, and high trade industries. Definitions are provided in section (5.2) of the text. Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies.

Table 7: Breakdown by Industry Group

Industry Group	OLS	IV1	IV2	IV1&IV2
	(1)	(2)	(3)	(4)
Durable Manufacturing	3.272 (0.217)*	2.574 (0.432)*	3.262 (0.35)*	3.045 (0.347)*
Non-Durable Manufacturing	2.946 (0.274)*	2.123 (0.642)*	2.540 (0.635)*	2.383 (0.612)*
Construction	3.274 (0.258)*	3.950 (0.6)*	2.868 (0.502)*	3.357 (0.461)*
Transport and Utilities	2.294 (0.245)*	2.396 (0.474)*	2.449 (0.42)*	2.449 (0.413)*
Wholesale and Retail Trade	3.063 (0.261)*	3.382 (0.524)*	3.215 (0.517)*	3.283 (0.497)*
Professional	2.173 (0.193)*	2.846 (0.414)*	2.575 (0.39)*	2.683 (0.378)*
Agriculture and Mining	2.742 (0.426)*	2.676 (0.817)*	2.464 (0.793)*	2.674 (0.81)*
Finance, Insurance, Real Estate	2.399 (0.31)*	3.966 (0.72)*	4.064 (0.688)*	4.033 (0.669)*
Public Administration	1.217 (0.221)*	1.692 (0.516)*	1.845 (0.467)*	1.782 (0.461)*
Others (Private, Business, Entertainment)	3.279 (0.331)*	4.115 (0.757)*	3.735 (0.678)*	3.856 (0.668)*
Obs.	28426	28426	28426	28426
R^2	0.564			

Notes: Each column is an estimate of equation (21) estimated separately by industry group. Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies.

Table 8: Breakdown by Education and Potential Experience

Experience	Ed. Group	OLS	IV1	IV2	IV1&IV2
		(1)	(2)	(3)	(4)
< 10	HS	2.706 (0.176)*	2.424 (0.327)*	2.997 (0.35)*	2.708 (0.245)*
	SP	2.394 (0.211)*	2.517 (0.518)*	3.029 (0.427)*	2.873 (0.389)*
	BA	1.780 (0.172)*	1.719 (0.645)*	2.323 (0.352)*	2.218 (0.343)*
> 10	HS	1.412 (0.172)*	3.959 (1.050)*	1.160 (0.367)*	1.602 (0.386)*
	SP	1.616 (0.357)*	2.014 (3.280)	1.879 (0.791)*	2.009 (0.743)*
	BA	0.993 (0.248)*	1.032 (1.616)	1.293 (0.385)*	1.253 (0.371)*
$\Delta ER_c \times \text{Ind.}$		Yes	Yes	Yes	Yes
Obs.		51535	51535	51535	51535
R^2		0.334			

Notes: Each column is an estimate of equation (21) stacked by education and potential experience group, where we have restricted common employment \times industry effects but is otherwise fully interacted. Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of industry-by-year dummies.

Table 9: Housing Price and Labor Force Growth

Variables	Log 1BR Rent			Log Labor Force		
	OLS	IV1	IV2	OLS	IV1	IV2
	(1)	(2)	(3)	(4)	(5)	(6)
ΔR_{ct}	3.505 (0.639)*	4.444 (1.024)*	3.103 (0.972)*	2.399 (0.772)*	6.198 (1.751)*	5.989 (1.658)*
ΔER_{ct}	1.633 (0.327)*	-.996 (1.878)	1.173 (1.234)	1.837 (0.476)*	12.919 (3.587)*	13.255 (2.988)*
Obs.	456	456	456	456	456	456
R^2	0.172			0.205		

Notes: Stars (*) denote significance at the 5% level. Standard errors, in parentheses, are clustered at the year-city level. All regressions contain a full set of year dummies.

Figure 1:

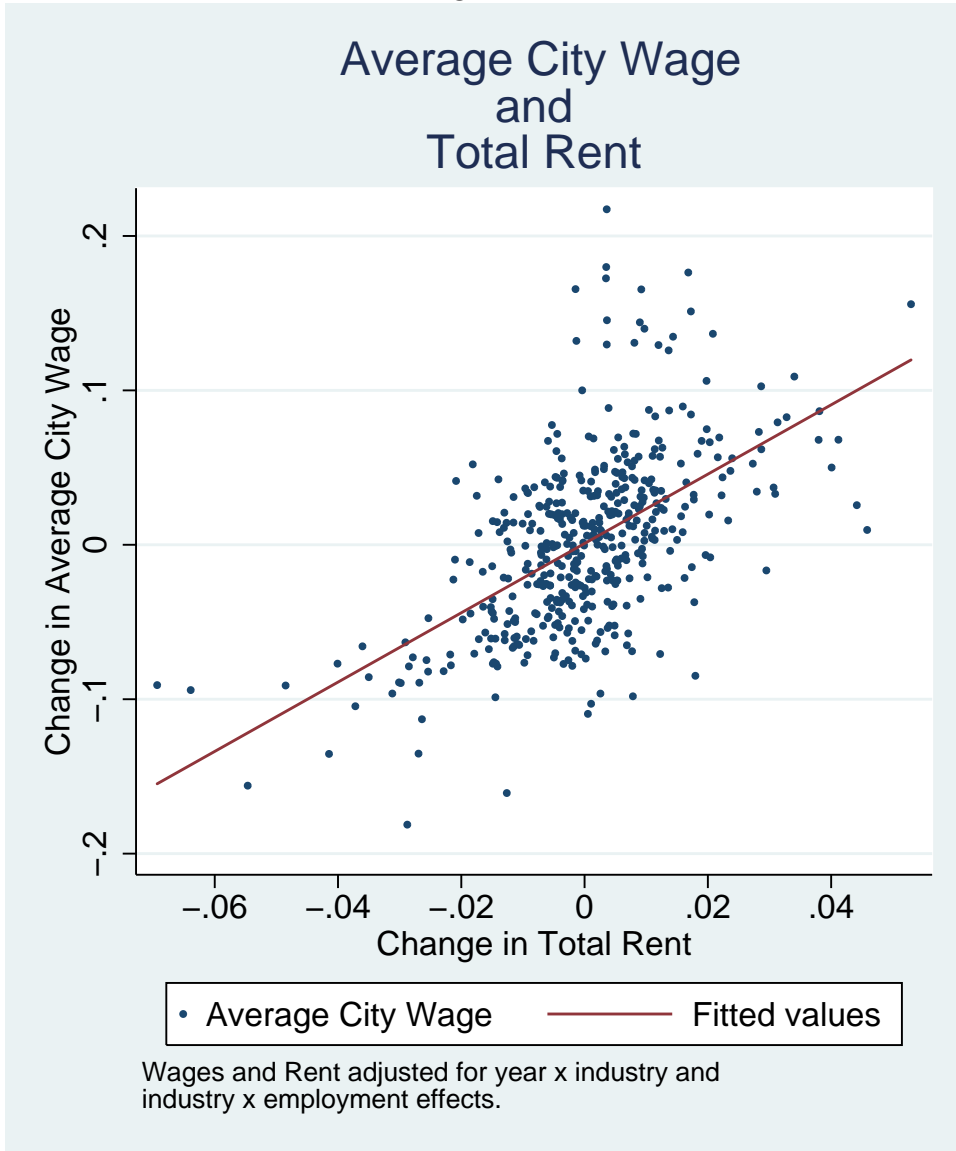


Figure 2:

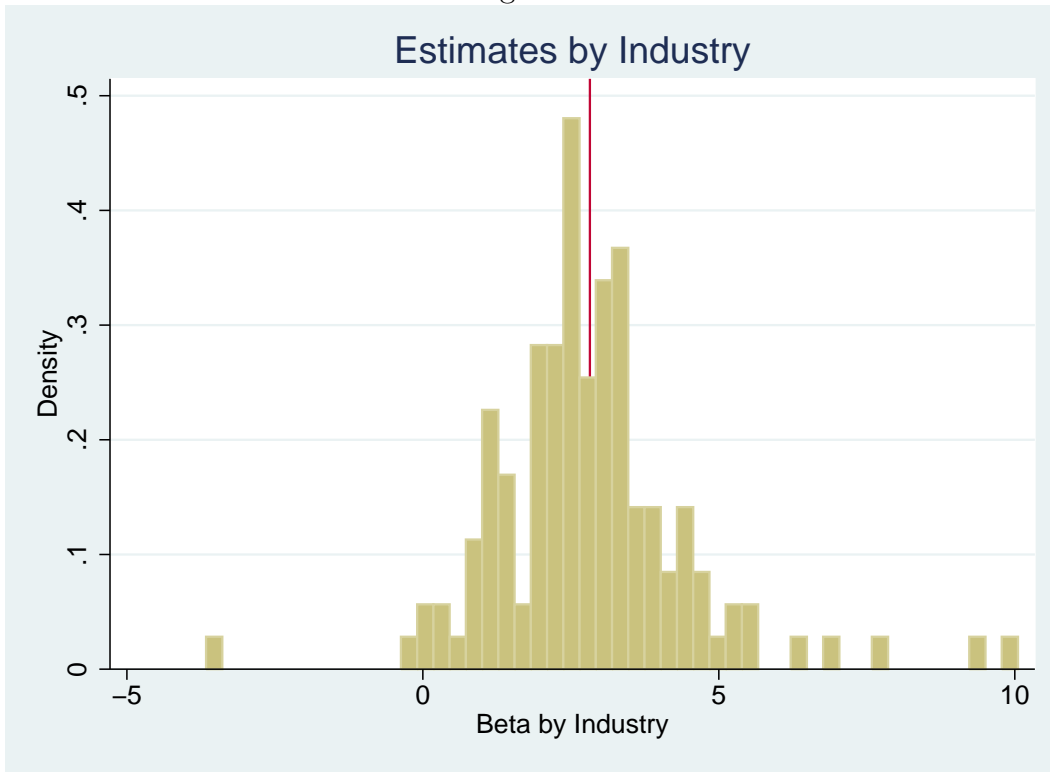


Figure 3:

