How Elastic is the Job Creation Curve?*

Paul Beaudry[†]
David A. Green[‡]
Ben Sand[§]

Februaryl, 2013

Abstract

We use search and bargaining theory to develop an empirically tractable specification of the job creation curve and to derive an instrumental variable strategy to estimate and test its validity. We estimate the job creation curve using city-level observations for 1970-2007. We find that U.S. city-industry level labor market outcomes conform well to restrictions implied by search and bargaining theory. Using 10-year differences, we estimate the elasticity of the job creation curve with respect to wages to be -0.3. We interpret this relatively low elasticity as reflecting a low propensity for individuals to become entrepreneurs when labor costs decline.

*We thank Chris Bidner, David Card, Giovanni Gallipoli, Patrick Kline, Alex Lefter, Kevin Milligan and seminar participants at Bocconi University, Cambridge University, LSE, McGill University, the University of California – Berkeley, University of Essex, University of Bristol and the University of Montreal for helpful comments and suggestions.

[†]University of British Columbia, and NBER

[‡]University of British Columbia and Research Fellow, IFS, London

[§]York University. bmsand@econ.yorku.ca

Introduction

The evaluation of policies or institutions that change labor costs requires a clear understanding of how such changes affect employment. Discussions of numerous policies, including payroll taxes, job security legislation, minimum wages and policies related to worker benefits are commonly framed in terms of whether they will "kill" jobs or, if cut, help create new ones.[1] As a recent example, in debates about stagnant employment outcomes as economies emerge from recession, many analysts have argued that cutting wage costs is the key to re-starting employment growth. Embedded in such a claim is the view that higher wages substantially reduce aggregate employment. The evidence on this, however, remains mixed. Many of the studies that examine the direct impacts of labour cost shifts – such as induced by payroll tax changes – find only modest impacts on aggregate employment (Blau and Kahn, 1999), implying rather inelastic labor demand curves. On the other hand, studies of regional responses to supply shocks tend to find quite small wage impacts, implying very elastic labor demand curves (Blanchard and Katz, 1992; Krueger and Pischke, 1997).[2] Further variation in estimates arises because different studies use different measures of employment outcomes (either employment rates or employment levels), often without a clear reference to theory to support their choice. Thus, policy makers face a confusing array of predictions about the employment impacts of policies that affect wage costs either directly or indirectly.

Attempts to provide estimates of the effects of wage changes on employment must deal with the inherent endogeneity problem stemming from the fact that wages and employment are jointly determined. In this paper, we provide estimates of the effects of wage costs on employment using a search and bargaining model to help guide and motivate our empirical strategy. We adopt this modelling perspective because it offers new insights regarding how to estimate the effects of wage costs on employment and it imposes testable implications which we can evaluate with available data. In addition, the model provides a potential explanation for disparities among earlier estimates of labor demand elasticities related to the use of different forms of variation and different dependent variables.

In search and bargaining models, labor demand is determined by an equilibrium relationship known as the job creation curve. This curve reflects employers' comparisons of the expected cost to the expected benefit of opening and maintaining a vacancy. With unrestricted entry, this implies that a rise in wages (which reduces the benefits of a filled job for the firm), must be offset by a decline in the tightness of the labor market, i.e., a decline in the employment rate. This equilibrium relationship is the job creation curve, and it is the relevance of this theoretical construct for

[1]For example, the California Chamber of Commerce releases an annual list of "job killer" bills which they claim identifies legislation that will "decimate economic and job growth in California". Often the identified bills contain "workplace mandates" which are argued to increase labor costs.

[2]The two extremes are captured in the minimum wage literature on one end (where studies commonly find either small positive or small negative elasticities) and the literature on city adjustments to shocks on the other (where, for example, Card (1990) finds virtually no wage response to the Mariel Boat-lift supply shock in Miami).

understanding employment determination we want to evaluate. The fact that the job creation curve relates wages to employment rates is important. A search and bargaining framework implies that employment rates and employment levels can react differently to wage changes and we suggest that this may explain the heterogeneity of elasticity estimates in the labour demand literature.[3]

In the canonical search and bargaining model the potential supply of job creators is assumed to be perfectly elastic. While this assumption greatly simplifies the calibration of job creation curves used in many macroeconomic models, it is not innocuous. Indeed, in many calibrated models the elasticity of job creation with respect to wages is very high precisely because the supply of job creators is assumed to be infinite. We do not want to impose this assumption, and instead follow Fonseca, Lopez-Garcia, and Pissarides (2001) in viewing job creators as being drawn from a pool of potential entrepreneurs with different abilities. The distribution of entrepreneurial abilities in the population then becomes a factor in determining the elasticity of job creation with respect to wages. If there is an abundance of good entrepreneurs, the job creation curve will be relativey elastic. On the other hand, if good entrepreneurs are scarce and are limited by span of control problems, then the job creation curve will be relatively inelastic. Our estimation of the job creation curve will allow us to assess this issue.[4]

Since we want to estimate the elasticity of job creation with respect to wages, we need to take a stance on the process of wage determination and its implications for admissible instruments. For this reason, we will also exploit properties of the second important locus in a search and bargaining model: the wage bargaining curve. This curve represents the outcome of bargaining over the match specific surplus between pairs of workers and employers. It is also a relationship between the wage and the employment rate since changes in labor market tightness alter the bargaining power of workers relative to employers. The equilibrium wage and employment rate are determined by the intersection of the job creation and wage bargaining curves. Thus, we can obtain estimates of the slope of the job creation curve if we can find exogenous shifters of the wage bargaining curve that are not directly related to the benefits of creating jobs.

Based on an earlier paper Beaudry, Green, and Sand (2012a, hereafter BGS), we argue that predictable shifts in the industrial composition of a local economy (which we will refer to as a city) can act as an instrument that moves the wage bargaining curve but not the job creation curve. In BGS we construct a multi-sector, multi-city version of a standard search and bargaining model and examine the impact of changes in the industrial composition of a city on wages within all industries in the city. The main idea is that in a search and bargaining model with multiple

[3] For example, within our search and bargaining framework a shift in labor costs (such as arises with changes in payroll taxes) can trace out the slope of the job creation curve, while a shift in the supply of available workers does not.

[4] The selection mechanism for entrepreneur heterogeneity we adopt shares similarities to other work that has modeled industries as collections of heterogeneous producers (Hamermesh (1993, Chapter 4), Jovanovic (1982), Melitz (2003)). The idea that there is heterogeneity among firms has received support from the related empirical literature (see, for example, Bartelsman and Doms (2000)).

sectors, a shift in the sectoral composition of the workforce toward relatively high-paying sectors improves the bargaining position (and thus the wages) of all workers, even holding labor market tightness constant. This arises because a worker's outside option in bargaining with his or her employer is based on the wages the worker can obtain in alternative employment options. However, in that paper we did not derive a specification for the Job Creation curve or estimate its slope, which is the focus of this paper. In this sense, the results in BGS serve to justify the first-stage regressions underlying the main analysis in this paper.[5] We refer readers to BGS for detailed derivations of parts of the model that are not directly related to the Job Creation curve.

The empirical approach we use in this paper is a structural-IV approach. By this we mean that we use a structural model to carefully derive the conditions under which proposed instruments provide consistent estimates of the coefficients of interest. However, we do not directly impose the model structure on the data. Instead, we estimate relatively straightforward linear regressions using the controls and instruments indicated from the theory and then test the over-identifying restrictions that are implied by the theory. Our goal is to allow the reader to fully understand the source of variation we are using to identify parameters and its relation to the theory. We depart from "full" structural estimation in that we focus on estimating only first order implications of the theory as implied by its linear approximation. We believe such a choice is desirable as it offers a simple, clear and intuitive exposition of results.

Our empirical work is based on U.S. Census data from 1970-2000 and data from the American Community Survey for 2007. Our approach relies on comparing industry-city level changes in employment rates between localities with different levels of wage pressure. The differential wage pressures arise from identified shifts in the bargaining position of workers induced by nation-wide shifts in industrial wage premia and composition. We look at effects over periods of 10 years (except for the shorter 7 year period 2000-2007), and therefore the estimates we find are associated with quite long run phenomena.

The main finding of the paper is that the type of labor demand specification implied by our augmented model of search and bargaining – which emphasizes employment rate-wage trade-offs – is given substantial support in the data. In particular, the specification and over-identifying restrictions implied by theory are easily accepted. Given this, we view estimates derived from this model as a reliable basis for assessing the impact of wage changes on employment outcomes. We present estimates of two types of elasticities identified by the model. At the city level, we find an estimate of -0.3, suggesting that the employment rate – labor cost trade-off is relatively inelastic. At the industry level, we find a larger elasticity, as predicted by the theory. We interpret the rather low city-level elasticity as partly reflecting a low propensity for individuals to become entrepreneurs and create jobs when wages are

[5] Working from the intuition in the model, we use a variant of Bartik (1993)'s approach to instrument for local composition shifts using national-level changes interacted with the start of period composition of the local workforce.

reduced. The model also provides a rationale for why studies focusing on regional adjustment yield results that seem to imply a very elastic labor demand curve. In particular, given a constant returns to scale matching technology and the assumption that potential job creators are proportional to the population, an exogenous inflow of workers simply replicates the economy, with no impacts on wages or employment rates. We show that the data conforms to this property, which fits with other investigations of the nature of the matching function (Blanchard and Diamond, 1989). Expressed in wage/employment-level space this means that an inflow of workers traces out a flat relationship as employment expands with no effect on wages. Our interpretation implies that this flat relationship is not related to a perfectly elastic labor demand, but is instead a series of equilibrium points reflecting adjustments on both sides of the labor market.

The remaining sections of the paper are as follows. In section 1 we propose an extended search and bargaining model and illustrate the implications of such a model for an empirical specification of a job creation curve. In particular, we use the model to derive both an empirical specification and an instrumental variable strategy for identifying key parameters. In section 2, we present our data and discuss implementation issues. In section 3, we present our main empirical results. The model we derive implies a number of estimating strategies and we illustrate how the results from each reinforces the other by satisfying various constraints that are implied by theory. In section 4 we assess the robustness of our results. Section 5 concludes.

Related Literature

A number of stands of literature are related to this paper. First, we build on standard equilibrium search and bargaining models (Pissarides, 2000) and extend the baseline framework to include endogenous industrial composition (Acemoglu, 1999, 2001), as industrial composition plays an important role in our identification strategy. Our primary goal is to use the search and bargaining framework to see whether it can shed new light on the link between wages and employment outcomes. We are unaware of other attempts to empirically evaluate the relevance of the job creation curve and estimate its slope. In standard expositions of Mortensen-Pissarides models, researchers generally assume a perfectly elastic supply of new vacancies, implying that the slope of the job creation curve depends solely on the rate at which vacancies are filled. Since the U.S. data indicates that vacancies are filled quickly (see for example Rogerson, Shimer, and Wright (2005)), this implies a calibration of the job creation curve which is very elastic with respect to changes in wages. However, as noted previously, we do not want to impose the assumption of a perfectly elastic supply of entrepreneurs but instead want to let the data speak. Accordingly in deriving our empirical specification of the job creation curve we follow Fonseca, Lopez-Garcia, and Pissarides (2001) in using a variant of the standard model in which heterogeneity in entrepreneurial ability allows for a potentially less than perfectly elastic supply of new vacancies. In this case the slope of the job creation curve will depend both on the speed at which vacancies are filled and on the intensity with which people will choose to become job creators as wages change.[6]

Our paper is most closely related to work directly investigating the impact of labor costs on employment, as surveyed by Blau and Kahn (1999), and to the large literature examining local adjustment to supply shocks (see, for example, Lewis (2004), Boustan, Fishback, and Kantor (2007) and Card (2009), among others). We depart from these studies by examining whether a search and bargaining perspective can provide a better understanding of the wage-employment nexus. Our focus on long run wage effects on employment differentiates our work from studies of regional adjustment to aggregate labor demand changes (Blanchard and Katz, 1992; Bartik, 1993, 2006) which mainly focus on unemployment dynamics.

Our identification of wage effects on employment uses variation in workers' outside options. This idea has precedents in the literature examining union wage and employment contracts (e.g., Brown and Ashenfelter (1986); MaCurdy and Pencavel (1986); Card (2009)) as these papers exploit measures of alternative wages outside the specific contract in their estimation. Card (1990) finds that the real wage in manufacturing has a positive effect on wage changes in the Canadian union contracts he studies, which echoes the mechanism underlying our basic source of identification. In a similar spirit, MaCurdy and Pencavel (1986) obtain estimates of production function parameters from data on wage and employment setting for typesetters when allowing for an alternative wage to effect the efficient outcome through an impact on union preferences.

In principle, our estimates should be easy to compare with those reported in the literature aimed at estimating the elasticity of labour demand derived within neoclassical models. Hamermesh (1993) provides the most comprehensive summary of the traditional literature estimating this elasticity. However, the question we ask is very different from the question in much of the traditional labour demand literature. We are interested in asking how employment changes when the cost of labour changes, knowing this will also affect the amount produced. In the labour demand literature the more common question is to ask how employment changes when wages changes holding constant total output. In fact, in what Hamermesh calls the "long run" elasticity, these elasticities are obtained from estimates of labour demand equations with both output and the relevant wage as regressors. Such estimates provide interesting information on the elasticity of substitution between capital and labour, but they are not directly comparable with the estimates we report here.[7]

Another difference between our framework and that of the traditional labour demand literature is that the labour demand literature generally uses either some measure of employment or the log of employment as the dependent variable. In contrast, we argue that a search model implies that the employment rate is the relevant dependent variable and that the distinction between employment levels and

[6] An important assumption we make regarding potential job creators is that their supply is proportional to the population. Our empirical results suggest that this proportionality assumption is a reasonable approximation.

[7] In Hamermesh (1993), the main estimates he reports lie in a range near -0.3, which suggest a rather low elasticity of substitution between capital and labour. While this elasticity is numerically very close to the one we obtain here, it is not appropriate to compare them as they do not address the same issue.

employment rates is important for understanding the different results in the literature regarding how employment responds to immigration relative to the literature on how it responds to an increase in the cost of labour.[8] Finally, the vast literature on estimating adjustment costs in a dynamic labour demand context (e.g., Bertola (1992)) is potentially relevant for our analysis given that the estimated adjustment times may be a reflection of the type of frictions that underlie search models. However the wage elasticity estimates from the dynamic labour demand literature are generally no more comparable to our estimates than those from the static demand literature.

1 Theoretical Framework

Our goal in this section is to derive an empirical specification for the job creation curve and show how such a job creation curve can be estimated with city-level data. To do this, we will work with the extended version of a standard search and bargaining model, similar to that presented in BGS, which incorporates multiple sectors, multiple cities, and endogenous entrepreneurial decisions. Here we focus on elements related to the job creation curve, in contrast to BGS which emphasized the process of wage determination.[9] Given our empirical focus, our model is highly stylized, but we will show that it still yields strong testable implications which will allow us to evaluate whether such a simple model is a reasonable approximation to the determination of employment.

At the heart of the model are search frictions that characterize the labor markets in all cities. For ease of presentation we will begin by assuming that workers are not mobile across cities. However, as we will see in a later section, the specification we derive here will be robust to including worker mobility and allowing expected utility to be equalized across cities.

Each local economy unfolds in continuous time and consists of firms and workers who are risk neutral, infinitely lived, and discount the future at a rate, r. Workers, for the moment, will be taken to be homogeneous in terms of skills. Firms and workers come together in pairs according to a matching technology and matches end at an exogenous rate, δ. Define L_c as the total available number of workers in city c, and E_{ic} and N_{ic} as the number of employed workers (or matches), and the number of available jobs in industry i in city c, respectively. Let $E_c = \sum_i E_{ic}$ and $N_c = \sum_i N_{ic}$. The number of matches in a city produced per unit of time is governed by the match-

[8] It has been pointed out to us that one could generate a specification with the employment rate as the dependent variable within a neo-classical framework if one assumes, as we do, that the number of potential entrepreneurs is proportional to the labour force and is an important factor of production. However, none of the papers directly estimating labour demand appear to have taken this approach. It is not our claim that it is impossible to derive our specification from a neoclassical model; only that the search model provides a direct rationale for the specification and implies an identification strategy and test that do not seem to us to arise naturally in other models.

[9] We refer readers to BGS for a more complete exposition of the model. However, since BGS does not contain several elements we use here, we also present details of the current model in Appendix A.

ing function:

$$M = M\left((L_c - E_c), (N_c - E_c)\right), \tag{1}$$

where the inputs are the available pool of unemployed workers, $L_c - E_c$, and the number of vacancies, $N_c - E_c$. As is standard in the search and bargaining literature, we assume the matching technology exhibits constant returns to scale and is increasing in both arguments. We will show that implications of this constant returns to scale assumption are supported by our data in the empirical section.

These properties of the matching technology imply that we can write the probability that a worker encounters a vacancy and the probability a firm fills a vacancy as:

$$\psi_c = \frac{M\left((L_c - E_c), (N_c - E_c)\right)}{L_c - E_c} \quad \text{and} \quad \phi_c = \frac{M\left((L_c - E_c), (N_c - E_c)\right)}{N_c - E_c}, \tag{2}$$

respectively. Using the steady state condition that the flow of workers leaving unemployment must equal the flow of workers exiting employment, these probabilities can be rewritten as functions of only the employment rate, $\frac{E_c}{L_c}$.[10] Thus, the matching probabilities are functions of the city-level employment rates. That is, we are assuming for simplicity that firms in all industries in a city are searching in the same labour pool.

1.1 Firm Bellman Equations

Firms in this economy will hire labor in a local labor pool, but will be assumed to be selling their output in the national market. The empirical specification for the job creation curve at the local level can be derived entirely from firms' Bellman equations and a specification on firm entry decisions. Denote by V^v_{ic} the present-discounted value of a vacancy in industry i and city c. In steady-state, which will be our focus, V^v_{ic} must satisfy the Bellman equation:

$$rV^v_{ic} = -h_i + \phi_c\left(V^f_{ic} - V^v_{ic}\right), \tag{3}$$

where h_i is the flow cost of maintaining the vacancy and with probability ϕ_c the vacancy is converted into a filled job, which has a present-discounted value of V^f_{ic}. In equilibrium, the latter value must satisfy

$$rV^f_{ic} = p_i - w_{ic} + \epsilon_{ic} + \delta\left(V^v_{ic} - V^f_{ic}\right), \tag{4}$$

where p_i is the national level price paid for the output of industry, i, w_{ic} is the wage paid to workers in industry i in city c, and ϵ_{ic} is an industry-city cost advantage where we assume $\sum_c \epsilon_{ic} = 0$. Equation (4) reflects the fact that firms in the same industry but different cities sell into the same goods market but can have different local cost advantages as caputured by ϵ_{ic}. As a result, once a match occurs, a firm enjoys a profit flow of $p_i - w_{ic} + \epsilon_{ic}$, and with probability δ the match is broken.

[10] In particular, substituting for $M(\cdot,\cdot)$ using the steady state condition, $\delta\frac{E_c}{L_c} = M\left(1 - \frac{E_c}{L_c}, \frac{N_c}{L_c} - \frac{E_c}{L_c}\right)$, we obtain $\psi_c = \frac{\delta\frac{E}{L}}{1-\frac{E}{L}}$.

Workers can either be employed or unemployed, and we assume that there are always gains from trade between workers and firms in all jobs created in equilibrium.[11] Once a match is made, workers and firms bargain a wage, which is set according to a Nash bargaining rule. We assume that workers only search while unemployed.[12]

The number of jobs created in industry i in city c, denoted N_{ic}, is determined by a free entry condition for entrepreneurs. Here we follow Fonseca, Lopez-Garcia, and Pissarides (2001) and view entrepreneurs as being drawn from the population and differing in terms of their capacity to manage many jobs. In particular, we assume each individual in city c (in addition to potentially being a worker) receives the option of becoming an entrepreneur in industry i with probability $\tilde{\Omega}_{ic}$. There is a fixed cost, K, to becoming an entrepreneur. On learning of the option to be an entrepreneur, the person also learns how many jobs he or she can manage, denoted by n, which is drawn from the distribution $F(n)$. Since the expected value of creating a job in industry i in city c is V_{ic}^{v}, all potential entrepreneurs with $n \geq \frac{K}{V_{ic}^{v}}$ will decide to pay the fixed cost and become an entrepreneur. Therefore, the number of jobs created in industry i in city c will be

$$N_{ic} = L_c \cdot \tilde{\Omega}_{ic} \cdot \int_{\frac{K}{V_{ic}^{v}}}^{\infty} n f(n) dn. \tag{5}$$

where, $\tilde{\Omega}_{ic}$ can be interpreted as a city-industry comparative advantage in creating certain types of jobs.

Our formulation of entry decisions implies that more efficient entrepreneurs enter the market and create jobs first. In fact, Equation (5) implicitly defines the number-of-jobs-to-population ratio in an industry-city cell, $\frac{N_{ic}}{L_c}$, as an increasing function of the value of a vacancy and as an increasing function of the comparative advantage term Ω_{ic}. For employment to expand, the value of creating jobs must rise in order to favor entry by the marginal entrepreneur who is of increasingly lower management capacity. This endogenous determination of entrepreneurs is an integral part of the mechanism by which changes in wages affect job creation in our model. One implication of having an imperfectly elastic supply of entrepreneurs for each industry-city pair is that cities will have diverse industrial structures. If we were to assume a perfectly elastic supply of entrepreneurs, the model would have the counterfactual implication that each city specializes in an inudstry.[13]

1.2 Determination of Employment

Using equations (3) and (4), we can write the value of a vacancy as:

$$rV_{ic}^{v} = \alpha_{c1} \cdot h_i + \alpha_{c2} \cdot (p_i - w_{ic} + \epsilon_{ic}), \tag{6}$$

[11] See BGS and Appendix A for statements of the Bellman equations related to the worker states.

[12] This is clearly a strong assumption. The assumption is not needed to derive our empirical specification for the job creation curve but it is useful for clarifying how we identify the wage effect on employment.

[13] In BGS, we use a somewhat different formulation for the fixed cost of opening a vacancy but the implications of the two formulations are the same. We adopt the one given in (5) because it is closer to Fonseca, Lopez-Garcia, and Pissarides (2001) and because it is intuitively more appealing.

where $\alpha_{c1} = -\frac{(r+\delta)}{r+\delta+\phi_c}$ and $\alpha_{c2} = \frac{\phi_c}{r+\delta+\phi_c}$. Note that both α_{c1} and α_{c2} depend on a city's vacancy contact rate, ϕ_c, which, in turn, depends on the city's employment rate, $\frac{E_c}{L_c}$. It is useful to make this relationship more explicit by taking a log linear approximation around the point where cities have an identical employment rate and industrial composition (this arises when cities have $\epsilon_{ic} = 0$, and $\tilde{\Omega}_{ic} = \Omega_i$, where Ω_i is an industry level job creation cost advantage that is common to all cities). In addition, to eliminate any fixed effects, it will be useful to work in changes in the log-linear approximation between two periods. Since the model is solved in steady state, we need to view the two time periods as two different steady states resulting from different draws of the exogenous shocks .[14] We can then rewrite (6) as:

$$\Delta \ln V^v_{ict} = \tilde{\alpha}_{it} - \alpha_2 \Delta \ln w_{ict} + \alpha_3 \Delta \ln \frac{E_{ct}}{L_{ct}} + \alpha_4 \Delta \epsilon_{ict}, \tag{7}$$

where $\Delta x_{it} = x_{it} - x_{it-1}$ and we have added time subscripts to correspond to the different periods (steady states).[15] The $\tilde{\alpha}_{it}$s correspond to industry-year specific effects, and, hence, can be captured in an empirical specification by industry-year dummy variables. The coefficients α_2 and α_3 (which are evaluated at the common vacancy filling rate that holds at the expansion point of the approximation) are predicted to be positive and negative, respectively.[16] This implies that the coefficients on both the industry-city specific wage and the city level employment rate should be negative since, all else equal, the return to opening new jobs will be lower when wages are higher and when the labor market is tighter (implying a longer time to fill a vacancy). The last term is an error term that depends on the city-industry cost advantages.

We now want to relate the change in the value of a vacancy to the employment rate in an industry. To do this, we first use the equilibrium condition that $\delta E_{ic} = \phi_c \cdot [N_{ic} - E_{ic}] \quad \forall i$ to replace $\frac{N_{ict}}{L_{ct}}$ with $\frac{E_{ict}}{L_{ct}}$ in (5) since we are interested in an expression in terms of employment rather than jobs. We then take a log linearization of this adjusted version of (5) and, taking differences between two periods, we arrive at:

$$\Delta \ln \frac{E_{ict}}{L_{ct}} = \theta_1 \Delta \ln V^v_{ict} + \theta_2 \Delta \tilde{\Omega}_{ict} \tag{8}$$

where θ_1 is positive since the incentive to create jobs increases as the value of vacancies increases. By replacing $\Delta \ln V^v_{ict}$ in (8) with its expression given in (7), we arrive at our main estimation equation - an expression for job creation in an industry city cell as a function of wages and overall labor market tightness:

$$\Delta \ln \frac{E_{ict}}{L_{ct}} = \varphi_{it} + \varphi_2 \Delta \ln w_{ict} + \varphi_3 \Delta \ln \frac{E_{ct}}{L_{ct}} + \zeta_{ict}, \tag{9}$$

[14]Differences between the steady states will arise from differences in local shocks, $\tilde{\Omega}_{ic}$ and ϵ_{ic}, and in national level, industry specific productivity.

[15]The explicit steps in the log-linearization are given in Appendix A.

[16]From the appendix, $\alpha_2 = \frac{\phi}{r+\delta+\phi}$ and $\alpha_3 = [\frac{r+\delta}{(r+\delta+\phi)^2} \cdot (h_i + p_i - w_i)] \cdot (\frac{\partial \phi}{\partial \frac{E}{L}}) \cdot \frac{E}{L}$, where ϕ is the value for the probability that a vacancy meets a worker that is common to all cities at the point around which the approximation is performed and $\frac{E}{L}$ is the common employment rate at the same point. α_3 is negative because $\frac{\partial \phi}{\partial \frac{E}{L}} < 0$, i.e., the tighter is the labour market, the lower is the probability a vacancy meets a worker.

where the φ_{it} represent a set of common industry-year effects, φ_2 is the elasticity of industry employment creation with respect to wages holding the aggregate employment rate constant, and $\zeta_{ic} = \theta_1\alpha_4\Delta\epsilon_{ic} + \theta_2\Delta\tilde{\Omega}_{ic}$. From the derivation in the appendix one can see that φ_2 is negative and is increasing (in absolute value) in θ_1. This arises because a high value of θ_1 means that the density function of management capacity is such that many jobs are created in response to a fall in wages. Hence, in this model, one of the reasons that the elasticity of employment with respect to wages can be low is that the supply of entrepreneurial activity may have a low elasticity with respect to variable profits. The more common assumption in the search and bargaining literature is that the cost of creating jobs is constant (often zero) implying that – holding aggregate tightness constant – job creation should be infinitely elastic. By estimating equation (9) we will be able to examine whether this assumption is appropriate.

The coefficient, φ_3 in (9) is predicted to be negative because a tighter labor market increases the duration of vacancies, thereby lowering the return to opening a vacancy in any sector. Thus, φ_3 captures classic congestion externalities in a search model extended to include multiple sectors. The implication that city-level changes in the employment rate have a negative effect on employment rates within industries is somewhat counterintuitive in the absence of a search model and, as a result, can be seen as an interesting testable implication of the framework. Note that φ_3 also increases in size as θ_1 increases, as a greater sensitivity of job creation to the value of a vacancy translates to a larger negative effect of employment rate on sectoral-level job creation.[17] The last term in (9), $\zeta_{ic} = \theta_1\alpha_4\Delta\epsilon_{ic} + \theta_2\Delta\tilde{\Omega}_{ic}$, is an error term that depends on changes in city-industry-level variable costs and job creation advantages.

Equation (9) forms one of the bases of our empirical investigation and can be interpreted as a job creation curve at the city-industry level. Notice that the estimation of this equation involves two challenges; a standard simultaneity problem between wages and employment, and a reflection problem (Manski, 1993; Moffitt, 2001) for the employment rates. To see the reflection problem, note that $\Delta \ln \frac{E_{ct}}{L_{ct}}$ can be approximated by $\sum_i \eta_{ict}\Delta \ln \frac{E_{ict}}{L_{ct}}$ so that (9) essentially involves a regression of the change in the log employment rate in a given city-industry cell, $\Delta \ln \frac{E_{ic}}{L_c}$, on its average across such cells within a city. We can address this issue via an instrumental variable strategy or alternatively solve out for $\Delta \ln \frac{E_{ct}}{L_{ct}}$. We will pursue both approaches to see whether they give similar results. However, this will not solve all of the simultaneity problems inherent in (9) since within-industry wage changes, $\Delta \ln w_{ict}$ are likely correlated with the error term in (9). Thus, to consistently estimate (9), we must also identify variation in wages that is uncorrelated with $\Delta\epsilon_{ic}$ and $\Delta\Omega_{ic}$. The following section discusses how to address the reflection problem and the subsequent section discuss aspects of the wage determination process for this search and bargaining framework which can be used to motivate relevant instrumental variables.

[17] Thus, tightness affects sectoral-level job creation indirectly though its effect on the value of vacancies. However, one could easily imagine that there are direct effects as well. For example, suppose we allowed potential entrepreneurs the possibility of being more attuned to entrepreneurial options when unemployment is high, by writing $\tilde{\Omega}_{ic} = \Omega_0 \cdot \frac{E_c}{L_c} + \Omega_i + \Omega_{ic}$, where $\Omega_0 < 0$. This would translate into a larger φ_3 coefficient, as it would reflect both types impacts of tightness on sectoral-level job creation. This would not alter our empirical framework, but would alter the interpretation of φ_3.

1.3 Deriving a city level job creation curve

Equation (9) is the job creation curve at the industry-city level. It will be informative to derive from it a job creation curve at the city level. To this end, let us first define η_{ict} as the fraction of employment in industry i in city c (i.e. $\eta_{ict} = \frac{E_{ict}}{\sum_j E_{jct}}$). Now consider aggregating Equation (9) using weights η_{ict}, and using the approximation $\sum_i \eta_{ict-1} \Delta \ln \frac{E_{ict}}{L_{ct}} \approx \Delta \ln \frac{E_{ct}}{L_{ct}}$, in order to get

$$\Delta \ln \frac{E_{ct}}{L_{ct}} = \frac{1}{1-\varphi_3} \sum_i \eta_{ict-1} \cdot \varphi_{it} + \frac{\varphi_2}{1-\varphi_3} \sum_i \eta_{ict-1} \cdot \Delta \ln w_{ict} + \sum_i \eta_{ict-1} \frac{\zeta_{ict}}{1-\varphi_3}. \quad (10)$$

This equation expresses the change in the employment rate within a city as being negatively affected by the average wage change in the city ($\sum_i \eta_{ict-1} \cdot \Delta \ln w_{ict}$), and positively affected by the weight sum of the φ_{it}. Notice that φ_{it} reflects a national-level effect associated with an industry. To express φ_{it} as a function of observables, we can rewrite (9) as

$$\Delta \ln E_{ict} = \varphi_{it} + \varphi_2 \Delta \ln w_{ict} + \varphi_3 \cdot \Delta \ln E_{ct} + (1-\varphi_3) \cdot \Delta \ln L_{ct} + \zeta_{ict},$$

and then take a average of this equation over cities (using the weights $\frac{1}{C}$, where C is the number of cities[18]). This gives

$$\sum_c \frac{1}{C} \Delta \ln E_{ict} = \varphi_{it} + \varphi_2 \sum_c \frac{1}{C} \Delta \ln w_{ict} + \varphi_3 \cdot \sum_c \frac{1}{C} \Delta \ln E_{ct} + (1-\varphi_3) \cdot \sum_c \frac{1}{C} \Delta \ln L_{ct},$$

where we have used the assumption that $\sum_c \frac{1}{C} \zeta_{ict} = 0$ since ζ_{ict} reflect relative advantages.

This equation implies that φ_{it} can be written as

$$\varphi_{it} = \sum_c \frac{1}{C} \Delta \ln E_{ict} - \varphi_2 \sum_c \frac{1}{C} \Delta \ln w_{ict} + d_t, \quad (11)$$

where d_t is a year effect that is common across cities. The first two terms on the right side of the above equation can be approximated as the growth of employment in industry i at the national level, denoted $\Delta \ln E_{it}$, and the growth of wages in industry i at the national level, denoted $\Delta \ln w_{it}$. Thus, equation (11) indicates that the industry specific intercept in 9) is approximately equal to the national level growth in employment in the industry corrected for the average wage growth in the industry. Using (11), we can write the job creation curve at the city level as

$$\Delta \ln \frac{E_{ct}}{L_{ct}} = d_t + \frac{1}{1-\varphi_3} \cdot \sum_i \eta_{ict-1} \cdot \Delta \ln E_{it} + \frac{\varphi_2}{1-\varphi_3} \cdot \sum_i \eta_{ict-1} (\Delta \ln w_{ict} - \Delta \ln w_{it}) + \tilde{\zeta}_{ct}, \quad (12)$$

where $\tilde{\zeta}_{ct}$ is the error term given by $\sum_i \eta_{ict-1} \frac{\zeta_{ict}}{1-\varphi_3}$.

Equation (12) now expresses cross-city differences in employment rate changes as a function of three main components. The first is a composition effect captured by $\sum_i \eta_{ict-1} \cdot \Delta \ln E_{it}$, which reflects that a city should have a better employment outcome if it is initially concentrated in industries which are growing quickly at

[18] Alternatively, we can use the weights $\omega_c = \frac{E_{c0}}{\sum_c E_{c0}}$, where 0 denotes an initial year

the national level. Second, we have a negative wage effect, which captures within industry adjustments to change in the cost of labor. This is given by the term $\sum_i \eta_{ict-1} (\Delta \ln w_{ict} - \Delta \ln w_{it})$, which is large if a city experiences wage growth across industries that is higher on average than that experienced nationally. Since φ_2 is negative, a high value of $\sum_i \eta_{ict-1} (\Delta \ln w_{ict} - \Delta \ln w_{it})$ will result in a worse outcome in terms of employment rates in the city. Finally, there is the error term that reflects changes in the city's comparative advantage.

A comparison of equations (12) and (9) reveals an important difference in the wage coefficients in each. The coefficient on the city-industry specific wage change in equation (9) is the direct effect of a wage change on the employment rate in an industry-city cell holding the aggregate employment changes in the city constant. This reflects the response of firms in an industry if that industry were too small to have a substantial effect on the overall equilibrium in the city. However, in general we would expect that the immediate effect of a wage change in i, as captured in φ_2, would only be a first-round response. The decrease in employment in i would imply a slacker overall labor market in the city which would raise the value of a vacancy for entrepreneurs to an extent captured by φ_3. Their resulting employment changes would then have further effects. The ultimate outcome of that process on total employment in the city would then be given by $\frac{\varphi_2}{1-\varphi_3}$, which is the coefficient on the aggregated wage change in (12). Given that φ_3 is predicted to be negative, the total impact of the wage change at the city level will be smaller than the direct, industry specific effect, reflecting the self-correcting nature of the congestion externalities.

Given this relationship between the two equations, we can view equation (12) in two ways: either as a regression of direct interest in its own right, allowing us to get an estimate of the full, equilibrium aggregate effect of wage changes; or as the first stage for the aggregate employment variable in (9), allowing us to get consistent estimates of the short-term, industry specific wage effects. Under either approach, the variable $\sum_i \eta_{ict-1} \cdot \Delta \ln E_{it}$ deserves special attention. We will call this variable, Z_1 . This type of variable was first discussed in Bartik (1993) and later used among others by Blanchard and Katz (1992), as an instrument to capture changes in local employment demand independent of supply factors. Interestingly, it appears naturally here as a shifter in the demand for employment.

When considering (12) directly, we are interested in the conditions under which Z_1 is uncorrelated with the error term. As a first step in this consideration, note that the error term, $\tilde{\zeta}_{ct}$, varies across cities while the aggregate employment changes, $\Delta \ln E_{it}$, that are used in Z_1 do not. This implies that aggregation from the city to the national level does not generate any type of endogeneity concerns when using Z_1 . This is a point that is common to the Bartik-type approach but is rarely recognized. On the other hand, potential correlation between the start of period industrial shares in a city (the η_{ict}'s), which are the source of the cross-city variation in Z_1 , and the error term is a potential concern. Recall that the error term consists of city level average changes in local productivity and vacancy cost shocks, both of which are defined as relative to national industry level shocks. High or low values of these local shocks will determine the industry composition of a city – the η_{ict}'s. Thus, one can show that the condition required for Z_1 to be uncorrelated with the error in (12) is a random walk type assumption under which the ϵ_{ict} and Ω_{ict} processes are such that η_{ict-1} is independent of $\Delta\epsilon_{ict}$ and $\Delta\Omega_{ict}$. Since the error term in (9) is also a function of ϵ_{ict} and Ω_{ict}, this same condition implies that Z_1 is a valid instrument for the aggregate employment change variable in estimating that equation. Thus, the key identifying assumption that we need is that changes in comparative cost advantages in a city are independent of their levels (and, thus, the city's industrial composition) at the start of the period. At first glance, this may seem to be a strong assumption, but we will show that it has associated with it a strong, testable over-identifying restriction that we can investigate to see if the assumption is valid. For now, we will maintain this assumption and turn to the other main identification concern: the endogeneity of the wage variables in (9) and 12). This is the goal of the next section.

1.3.1 Instruments for Wages

To estimate the parameters in (9) and (12) consistently, we need to isolate forces that change wages in a city but that are not related to the change in comparative advantage terms that are in the errors of both equations. Given our search and bargaining framework, to isolate this type of variation, we need to find factors that change workers' bargaining power for given productivity levels. To do this, we will exploit insights presented in BGS regarding the role of industrial composition in affecting bargaining power and thereby wages. The idea in BGS is simple. Consider two identical workers bargaining their wages with employers in the same industry but different cities. If there are any frictions hindering perfect and costless mobility across cities then the outside options of these workers will differ since their fall back position will depend on the state of the labor market in their respective cities. In particular, if workers have some mobility across industries then workers' outside options, and hence their bargaining power, should be sensitive to the industrial composition of their cities, even if the tightness of the two labor markets is the same. Theoretically, workers in, say, chemical industry should be able to bargain a higher wage if they live in a city with high paying steel mills than if they live in a city where the steel mills are replaced with low paying textile mills. It is this idea we want to exploit to justify two instrumental variables that will allow us to consistently estimate (9) and (12). The two instruments will be valid under the same assumption that makes the Z_1 a valid instrument, that is, under the assumption that changes in a city's comparative advantage acts like a random walk and therefore is independent of past industrial composition.

To better understand this potential source of identification, it is helpful to look at the Bellman equations for employed and unemployed workers. If we let U^e_{ic} represent the value function of an employed worker in industry i in city c, then it should satisfy

$$\rho U^e_{ic} = w_{ic} + \delta(U^u_{ic} - U^e_{ic}), \tag{13}$$

where w_{ic} is the wage in industry i and city c, δ represents an exogenous separation rate and U^u_{ic} represents the value associated with being unemployed when the worker's previous job was in industry i.

When an individual is unemployed, assume that he gets flow utility from an unemployment benefit b and that unemployed individuals find jobs at rate ψ_c. What type of job will this individual find? It seems reasonable to believe that the individual is more likely to find a job in his previous industry, but it also seems reasonable to believe that the worker may – at least occasionally – find a job in another industry. Moreover, if he is to find a job in another industry, the identity of that industry is likely positively related to the local prevalence of that industry. To capture these ideas, assume an unemployed worker locates a job with probability ψ_c regardless of his former industry of employment. Conditional on finding a job, he finds a job in his former industry, i, with probability μ, and with probability (1-μ) he gets a random draw from jobs in all industries (including i). In this case, the value associated with being unemployed satisfies

$$\rho U_{ic}^{u} = b + \psi_c \cdot \left(\mu U_{ic}^{e} + (1-\mu) \sum_j \eta_{jc} U_{jc}^{e} - U_{ic}^{u} \right). \tag{14}$$

Using these two equations, we can represent the gain from being employed relative to being unemployed as,

$$U_{ic}^{e} - U_{ic}^{u} = \frac{w_{ic} - b}{\rho + \delta + \psi_c \mu} - \frac{\psi_c (1-\mu) \sum_j \eta_{jc} (w_{jc} - b)}{(\rho + \delta + \psi_c \mu)(\rho + \delta + \psi_c)}. \tag{15}$$

Assuming that wages are set to split the surplus between workers and employers – which is the common assumption in the search and bargaining literature – a force that reduces this difference will imply a higher bargained wage. Such a force is found in the $\sum_j \eta_{jc} w_{jc}$ element in equation (15). Equation (15) shows that, within a given industry, bargaining power (and, hence, wages) will be higher in cities with a higher overall average wage. Notice that this is not a mechanical result since the ability of workers to switch industries implies that it would arise even if we just focused on other industries by dropping i from $\sum_j \eta_{jc} w_{jc}$.

It is useful to decompose the movements in $\sum_j \eta_{jc} w_{jc}$ as follows:

$$\Delta \sum_j \eta_{jct} w_{jct} \approx \left(\sum_j \eta_{jct-1} (w_{jct} - w_{jct-1}) \right) + \left(\sum_j w_{jct-1} (\eta_{ict} - \eta_{ict-1} \right). \tag{16}$$

Equation (16) indicates that for a worker in a particular city, his outside option (and bargaining power) will increase over time if he is in a city where either wages within industries increase on average or where there is a shift in industrial composition toward relatively high-paying jobs. Each of these components forms the basis for an instrument for wages in (9) and (12). For the first instrument, consider the term $\sum_j \eta_{jct-1}(w_{jct} - w_{jct-1})$ in (16). Clearly, it cannot be used as an instrument as it stands because the local wages used in its construction will be correlated with the local productivity terms in the regression error terms. But we can obtain a potentially useful instrument by replacing the wages in this expression – which are at the industry city level – with wages at the national level as given below:

$$Z_2 = \sum_j \eta_{jct-1} (\ln w_{jt} - \ln w_{jt-1}).$$

This new variable, which we will call Z_2, is one of our candidate instruments for the wages changes in (9) or (12). For it be a valid instrument, Z_2 must have predictive power for wage changes at the industry-city level and needs to be uncorrelated with the errors terms of these two equations. We refer the reader to BGS for a formal exposition that Z_2 should be a good predictor of wage growth at the industry-city level, but our previous discussion about the sources of bargaining power should make this intuitively clear: the bargaining power of workers in a particular city should increase if the city is concentrated in industries which are experiencing high wage growth at the national level. To determine whether Z_2 is correlated with the error terms, note that our estimating variation comes from cross-city variation and that this variable gets its cross-city variation entirely from the η_{ict}'s (the local industrial composition). As in our discussion of Z_1, the national-level wage changes are not relevant for our consistency considerations since they are common across cities. Also, as with Z_1, this implies that Z_2 will be uncorrelated with the error terms in (9) and (12) under the assumption that the comparative advantage terms ϵ_{ict} and Ω_{ict} behave as random walks with changes independent of past levels.[19]

The second instrument for wages we want to propose builds on the second term of (16), that is, on $(\sum_j w_{jct-1}(\eta_{ict} - \eta_{ict-1})$. As in the previous case, this term would not be an appropriate instrument as it will not be orthogonal to the error terms in (9) or(12). Thus, consider a the closely related variable given by

$$Z_3 = \sum_i \ln w_{it-1} \cdot (\hat{\eta}_{ict} - \eta_{ict-1}) = \sum_i \eta_{ict-1} \cdot (g^*_{it} - 1) \cdot \ln w_{it-1}, \tag{17}$$

where $g^*_{it} = \frac{1+\Delta \ln E_{it}}{\sum_j \eta_{jct-1}(1+\ln \Delta \ln E_{it})}$. For the variable Z_3, we have again replaced wages at the city level by their national-level counterpart and we have replaced the industrial composition term η_{ict} with it predicted value base on η_{ict-1} and national-level trend in employment patterns.[20] As with Z_1 and Z_2, the resulting variable gets its cross-city variation from the η_{ict}'s and the same random walk assumption is needed for consistency. Furthermore, it should have predictive power for industry-city wages changes as it it should capture the higher bargaining power of workers in a city where we predict that the industrial composition is tilting toward the highest paying jobs.

We have now presented three instruments that can be used to estimate (9) and (12): Z_1, Z_2 and Z_3, with all three valid under the same random walk assumption. While this is an assumption that is implicit in the literature (for example in Blan-

[19]BGS presents a formal derivation of the form of the error term in the wage equation and, working from that, prove that the conditions listed here imply that these instruments are valid.

[20] To create the predicted share term, we first predict the level of employment for industry i in city c in period t as:

$$\hat{E}_{ict} = E_{ict-1}\left(\frac{E_{it}}{E_{it-1}}\right).$$

Thus, we predict period t employment in industry i in city c using the employment in that industry-city cell in period $t-1$ multiplied by the national-level growth rate for the industry. We then use these predicted values to construct predicted industry-specific employment shares, $\hat{\eta}_{ict} = \frac{\hat{E}_{ict}}{\sum_i \hat{E}_{ict}}$, for the city in period t.

chard and Katz (1992)), it is clearly a strong assumption, and it would be desirable to test it. Interestingly, since we have two instruments for the wages in these equations, we have an over-identified problem which implies that the validity of the underlying identification assumption can be evaluated. In particular, if the identification assumption is valid, it suggests that we can estimate (9) using either Z_2 or Z_3 as an instrument for the wage changes and we should get similar estimates of the slope of the job creation curves. On the other hand, both instruments will be invalid if the η_{ict-1}'s are correlated with the $\Delta\epsilon_{ic}$'s and $\Delta\Omega_{ic}$'s in the error terms and these offending correlations will be weighted differently by the two instruments (with changes in national-level industrial wages in Z_2 and national-level employment changes in Z_3). This would, in turn, imply that the two instruments should result in quite different estimated coefficients. Further, we view the over-identification test as strong because Z_2 and Z_3 work from quite different variation. Z_2 emphasizes improvements in worker bargaining power because of changes in wages, holding industrial composition constant. On the other hand, Z_3 emphasizes improvements stemming from changes in industrial composition holding industrial wage premia constant. In our data, the two variables have a correlation of only 0.18 after removing year-effects (as we do in all of our estimations).

In BGS, we focus on working from (15) to derive a wage equation and analyze the impact of shifts in industrial composition on wages. Here, we will only present results from the reduced form derived from that structural equation since our focus is on estimation of the job creation curve. The reduced form wage equation is given by

$$\Delta \ln w_{ict} = d_{it} + \beta_2 Z_{2ct} + \beta_3 Z_{1ct} + \Delta\xi_{ict}, \tag{18}$$

where the d_{it}s are time-varying industry effects. We estimate variants of (18) where we included either Z_2 or Z_3 or Z_2 and Z_3 together. We refer to BGS for the derivation of the structural wage equation and the composition of the error term.

1.4 Worker Heterogeneity

As we have indicated, our aim in this paper is to provide an estimate of how employment, on average, is affected by an across-the-board increase in the cost of labor. By its very nature, this question is about an aggregate labor market outcome. In the model developed so far all workers are identical and so all parameters are "aggregate" by definition. However, in our data, workers are heterogeneous in, for example, education and experience. We need to address this heterogeneity in order to provide a consistent answer to our aggregate question. Depending on the assumptions that one makes, there are several ways to approach this issue.

Our first approach, which we use for our main set of results, is to treat individuals as representing different bundles of efficiency units of work, where these bundles are treated as perfect substitutes in production. Therefore, we control for skill differences in wages via a regression adjustment. However, the theory indicates that we should not account for differences in worker attributes when aggregating the number of workers across industries and cities in the calculation of employment rates. In

Appendix D, we present a formal justification for such an approach. Heuristically, the structure of the basic search model allows one to write the wage of any skill type in units of some arbitrary type, which rationalizes the use of the regression adjusted wages. In addition, random matching implies that the probability of meeting a worker of a given skill type will equal the product of the vacancy filling rate and the proportion of that worker's type in the local economy. Thus, the effects of labor market tightness and skill can be separated and, we argue in the appendix, that the latter effects can be expected to be relatively small. Hence, the value of a vacancy will depend on the skill distribution of the workforce in general, but it is probably only of second-order importance. Our approach is to ignore these additional effects in our baseline specification. However, in the robustness section we will show that our results are not sensitive to controlling for measures of the aggregate distribution of skill in a locality.

An alternative assumption is that labor markets are segregated along observable skill dimensions and that our model applies to homogeneous workers within these markets. Thus, we also perform our analysis separately by education group as a specification check. Finally, if one assumes that the matching function is not over workers but over efficiency units then it is appropriate when aggregating workers over industries and cities to use efficiency weighted worker counts. When we construct our employment variables in this manner it does not substantially affect our main results. However, we choose not to pursue this interpretation as it appears quite artificial in the search and bargaining context. Nonetheless, results using efficiency units are available upon request.

1.5 Mobility

The model presented above assumes that workers are not mobile across localities. It may seem, at first, that allowing for worker mobility could overturn the result that wages differ across localities because of local bargaining conditions. However, even when we allow for directed search across cities, wage differentials between cities will generally continue to persist. In Appendix E, we offer two extensions of the model that take into account worker mobility. In the first extension, unemployed workers are not perfectly mobile, but are only occasionally offered the opportunity to switch cities. Due to this friction, wages (and outside options) will maintain a local component. When the option to switch cities arises, workers choose the city that maximizes the value of their search. Since this choice will not depend on the initial location of the worker, it acts as a common element across workers and is captured by an intercept.[21] In this extension of the model, none of our empirical specifications are affected and the model's implications will continue to hold.

In the second extension, we go further by modelling the spacial equilibrium more explicitly. In particular, we introduce local housing prices and allow workers to choose a city that maximizes expected utility, taking into account housing costs and local amenities. In this extension, we modify workers' Bellman equations, (14) and

[21]Random search across cities has the same implication.

(13), to include a negative function of the local housing price. Importantly, housing prices will not directly affect wage negotiation because it is a cost that is incurred whether or not an individual is employed (and wages depend on the difference $U^e_{ic} - U^u_{ic}$). However, housing prices will have to adjust to equilibrate expected utility across cities. In Appendix E, we present this simple model with housing costs that depend positively on local amenities and city size. We show that it implies that a city with a higher employment rate or higher expected wages (i.e., due to an industrial composition that is biased towards high-paying industries) will attract more workers. This in-migration will drive up local housing costs to the point where expected utility is equalized across cities, but in-migration will stop before wage equalization occurs. This is good from the point of view of this paper since what we need to identify the JC curve is cross-city variation in wages.[22] Moreover, the forces we emphasize in our model also have implications for worker mobility and housing costs. We view these as important issues and we investigate them further in related work (Beaudry, Green, and Sand, 2012b).

2 Data Description and Implementation Issues

The data we use in this paper come from the U.S. decennial Censuses from the years 1970 to 2000 and from the American Community Survey (ACS) for 2007. For the 1970 Census data, we use both metro sample Forms 1 and 2 and adjust the weights for the fact that we combine two samples.[23] We focus on individuals residing in one of our 152 metropolitan areas at the time of the Census. Census definitions of metropolitan areas are not comparable over time. The definition of cities that we use in this paper attempts to maximize geographic consistency across Census years. Since most of our analysis takes place at the city-industry level, we also require a consistent definition of industry affiliation. Details on how we construct the industry and city definitions are left to Appendix C.

As discussed earlier, our approach to dealing with worker heterogeneity is to control for observed characteristics in a regression context. Since most of our analysis takes place at the city-industry level, we use a common two-step procedure. Specifically, using a national sample of individuals, we run regressions separately by year of log weekly wages on a vector of individual characteristics and a full set of city-by-industry dummy variables.[24] We then take the estimated coefficients on the city-by-industry dummies as our measure of city-industry average wages, eliminating all cells with fewer than 20 observations.

[22] Formally, for the wage variation remaining after mobility has equalized utility across cities to be useful, the costs of housing must enter the valuation functions of firms and workers differently. This seems to us to be a reasonable assumption.

[23] Our data was extracted from IPUMS, see Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander (2004)

[24] We take a flexible approach to specifying the first-stage regression. We include indicators for education (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.

Our interpretation of the regression adjusted wage measure is that it represents the wage paid to workers for a fixed set of skills. However, since we only observe the wage of a worker in city j if that worker chooses to live and work in j, self-selection of workers across cities may imply that average city wages are correlated with unobserved worker characteristics such as ability. In this case, our wage measure will not represent the wage paid per efficiency unit but will also reflect (unobservable) skill differences of workers across cities. To address this potential concern, when we estimate our wage equations we control for worker self-selection across cities with a procedure developed and implemented by Dahl (2002) in a closely related context.

Dahl proposes a two-step procedure in which one first estimates various location choice probabilities for individuals, given their characteristics such as birth state. In the second step, flexible functions of the estimated probabilities are included in the wage equation to control for the non-random location choice of workers.[25] The actual procedure that we use is an extension of Dahl's approach to account for the fact we are concerned with cities rather than states, as in his paper, and that we also include individuals who are foreign born. When we estimate the wage equations, the selection correction terms enter significantly, which suggests that there are selection effects. Our results with or without the Dahl procedure are very similar. Nevertheless, all estimates presented below include the selection corrected wages.[26]

One of our main covariates of interest 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 individual characteristics 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 industry premia in constructing our R measures.

The dependent variable in our analysis is the log change in industry-city-level employment rates. We construct the industry-level employment rate by summing the number of individuals working in that particular industry divided by the city working-age population. Using employment to population ratios means that we include individuals who are classified as being out of the labor force as being relevant for labor market tightness. This is consistent with previous work on matching functions (Blanchard and Diamond, 1989) and on local labor market conditions (Bartik, 2006). Nevertheless, we have assessed the sensitivity of our results to this assumption and found it to be robust to an alternative definition of employment rates that restricts the population to include only those individuals that report themselves as being in the labor force. 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-

[25]Since the number of cities is large, adding the selection probability for each choice is not practical. Therefore, Dahl (2002) suggests an index sufficiency assumption that allows for the inclusion of a smaller number of selection terms, such as the first-best or observed choice and the retention probability. This is the approach that we follow.

[26]Details on our implementation of the Dahl's procedure are contained in Appendix E. Results without the selection correction are very similar to those reported here, and are available upon request.

1990) plus the 2007-2000 difference, 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.[27]

3 Results

We begin the discussion of our results by presenting estimates of the first-stage, or reduced form, wage equation. These echo results presented in BGS where the focus is precisely on the estimation of the impact of industrial composition shifts on wages, though the specifications are not identical because our focus, here, is on the reduced form. [28]

The columns of Table 1 contain estimates for the first-stage wage equations. In Columns 1 and 2, we regress wage changes on Z_{2ct} and Z_{3ct}, respectively. This and all subsequent specifications include a full set of year-by-industry dummies (4×144), but we suppress the presentation of these coefficients for brevity. The wage equation derivation in BGS indicates that Z_1 should be included and so it is a covariate in all the wage regressions here. The key point from these columns is that both Z_{2ct} and Z_{3ct} are strong and highly significant determinants of $\Delta \ln w_{ict}$. Recall from our earlier discussion that Z_{2ct} and Z_{3ct} correspond to quite different parts of the variation in $\sum_j \eta_{jc} w_{jc}$ but the theory implies that they should have the same effect on bargaining power and, hence, on wages. The estimated coefficients on the two variables are very similar, and we cannot reject the restriction that they are equal at any conventional significance level. We view this as a strong piece of supportive evidence for the model and our identifying assumptions. In the last column of Table 1, we include both Z_2 and Z_3 to show that they are both individually helpful in predicting wages.

3.1 Estimates of the Job Creation Curve: Basic Results

In Table 2, we present estimates of our main equation of interest (9). As in Table 1, all the reported regressions include a full set of year-by-industry dummies. Column 1 reports OLS results. Both the coefficients on the wage and the city-level employment rate are positive and highly significant. This is the opposite of what our theory predicts for consistent estimates of the coefficients in (9). However, the employment rate equation derived from the model implies that OLS estimation of this equation should not provide consistent estimates. Moreover, the fact that productivity shocks, ϵ_{ict}, enter the wage and employment rate equation error terms suggests that the OLS regression coefficients are likely to be positive.[29]

[27]We cluster at the city-year level because this is the level of variation in our data. Clustering only by city has little effect on the estimates of standard errors that we report.

[28] BGS focus on investigating and interpreting the impact of changes in industrial composition on $\Delta \ln w_{ict}$. In BGS it is argued that the estimates imply that the impact of a change in industrial composition on the average wage in a city is approximately 3.5 times what is implied by a standard decomposition approach once spill-over effects from bargaining are taken into account.

[29]See BGS for the derivation of the composition of the error term in the wage equation.

Columns 2-4 contain results from estimates of (9) using the Z_2 and Z_3 instruments discussed in section 1.3.1. As we saw in Table 1, the instruments perform well in the first-stage. F-statistics for wages are reported in the bottom rows of Table 2, supporting our claim that our instruments are good predictors of within-industry wage changes and we do not face weak instrument concerns. As we will see in the aggregate equation results presented in the next table, Z_1 also performs well in the first stage for the city level employment rate.

The IV results indicate that labor costs are negatively associated with sectoral employment rates, as predicted by theory. The magnitude of the estimated coefficients on $\Delta \ln w_{ict}$ obtained from either Z_2 or Z_3 are nearly identical, and a standard over-id test fails to reject the null hypothesis of equality at any conventional significance level.[30] As argued earlier, we view this result as being particularly important. From a theoretical point of view, bargaining implies that improvements in workers' outside employment opportunities should have the same impact on wages regardless of whether the improvements arise from growth in a high paying industry in a locality (the variation emphasized in Z_{3ct}) or increases in industrial premia in existing industries (the variation emphasized in Z_{2ct}). Given this, we also expect labor demand responses estimated from either source of variation to be the same since employers are only concerned about the bargained wage. From an empirical point of view, this result provides an important test of our identifying assumptions.

The other key prediction from the model is that an increase in labor market tightness in a city (as represented by the city-level employment rate) should negatively affect within industry employment rates. Once we instrument, we do, in fact, find a strongly significant negative effect of $\Delta \ln \frac{E_{ct}}{L_{ct}}$ on $\Delta \ln \frac{E_{ict}}{L_{ct}}$. This is a striking result since one may expect a positive relationship between these variables. As we discussed earlier, it is difficult to reconcile this result with a standard neoclassical model but it is a clear prediction of a search and bargaining framework. Recall that the size of both the coefficient on wages and the coefficient on the aggregate employment rate are affected by θ_1, the elasticity of job creation with respect to the value of a vacancy. If this elasticity were close to infinity, as is generally assumed in most search and bargaining models, we would expect both these coefficients to be close to infinity. While these coefficients are quite large, neither are close to infinity, suggesting that job creation is far from being perfectly elastic with respect to the value of vacancies.

We are now in a position to interpret the results. Consider a wage increase in a particular industry, holding overall employment rates constant. If the industry in question is not so large as to have a significant impact on overall employment rates, the estimates using Z_2 and Z_3 as instruments in Table 2 imply a labor demand elasticity at the industry level of about -1.

What about improvements of wages in a city more generally? Since all industries will adjust employment downward in response to a general wage increase, there will be feedback effects on overall employment rates. Allowing these equilibrium effects to play out implies a city-level labor demand elasticity of $\frac{\varphi_2}{1-\varphi_3}$ or of about -0.30 given

[30]The actual test statistic value is 0.17 with a p-value of 0.68.

our estimates. In other words, since φ_3 is predicted to be less than zero in a search and bargaining model, overall wage increases in a locality have a built in dampening effect on employment responses because they simultaneously increase the availability of workers. In our model, this leads to reduced search costs for firms (because vacancy contact rates are higher) or improved average entrepreneur quality. Thus, our estimates suggest that city-level job creation curves are relatively wage inelastic, with an increase in wages of 10% impling a reduction in the city-level employment rate of about 3%.

Recall that we can also obtain an estimate of the city-level demand elasticity through direct estimation of our city-level regression, (12). Estimates of (12) are presented in Table 3. This table has a similar format to those that proceed it, and also contains a full set of year dummies. IV estimates of the coefficient $\frac{\varphi_2}{1-\varphi_3}$ in columns (2)-(4) range from about -0.27 to -0.31. The estimates obtained using Z_2 and Z_3 are again nearly identical and a standard over-identification test again fails to reject the null hypothesis. Thus, in this specification, too, the results strongly support the search and bargaining model. Furthermore, the estimates of the city-level demand elasticities using the aggregated data are almost identical to what we just calculated using the estimated coefficients from specification (9). Since estimation of (9) and (12) use very different levels of variation (and since there is no mechanical reason the two specifications should provide the same results), we view the similarity of the estimates of the city-level elasticity obtained from each as a further piece of evidence supporting the search and bargaining model for understanding the wage-employment nexus. Finally, note that Z_1 has a positive and strongly statistically significant direct effect on the city level employment rate, implying that it is a strong instrument for that employment rate in the disaggregated equation estimation.

4 Robustness and Specification Tests

In this section, we assess the sensitivity of our results to a variety of specification checks.

4.1 Exploring the Effects of Labor Force Size on Employment Determination

Table 4 contains results from our baseline specification (9) and our aggregated specification (12) where we include labor force growth as an additional control variable. Under the assumptions of the model (specifically, that the matching function exhibits constant returns to scale and that the number of entrepreneurs is proportional to the labor force), changes in the size of the labor force should have no impact on employment rates once we control for local demand conditions and labor costs. The estimated coefficient on labor force growth in these equations serves as a check on these assumptions. The first column shows OLS estimates, while in Columns 2-4 we instrument for changes in within-industry wage and city employment rates. In columns 1-4, we do not instrument the city's labor force variable. Interestingly, when we instrument the wages and employment rate when including the labor force growth variable, the estimated coefficients on $\Delta \ln w_{ict}$ and $\Delta \ln \frac{E_{ct}}{L_{ct}}$ change very little from the results presented in Table 2. It should be kept in mind that since we are examining decadal differences, these results should be interpreted as the long-term consequences of increasing the size of a city's labor force.

The growth in labor force is not likely to be exogenous. We attempt to address this possibility by constructing an instrument set for labor force growth that is based on long-term city climate variables. Since mobility is likely driven by local amenities, variation in local amenities would potentially provide a good set of instruments. However, most amenities (not related to employment and wages) are relatively constant over time, making them unhelpful as instruments in our difference specification. Nonetheless, measures of amenities can still be used as instruments in this case if the value of the amenity has changed over time. For example, if the value of living in a nice climate has increased over time then the level of the climate indicator variable can be used as an instrument variable for labor force growth.[31] Building on this insight, we collected data from a number of sources to construct an instrument set consisting of average temperatures and precipitation for each city in our sample. Consistent with the idea that workers are increasingly drawn to cities by amenity factors, we find that indicators of mild climates are significant predictors of city labor force growth.[32]

The estimates in Column 5 of Table 4 show the results using all of the instruments including the climate variables. The first stage F-statistic on labor force growth, reported in the second last row of the table, is 24.0, indicating that we predict labor force growth quite well. The estimated coefficient (standard error) on the labor force growth variable is -0.12 (0.072), which is larger than the non-instrument results but still not statistically significant. The estimated coefficients on $\Delta \ln w_{ict}$ and $\Delta \ln \frac{E_{ict}}{L_{ct}}$, are slightly lower, but still imply a city labor demand elasticity of $\frac{-.77}{1+1.28} = -0.34$, which is in line with our previous estimates.

In Column 6, we report results from the city-level specification (12) with labor force growth included and using the climate instruments. This result supports the idea that a job creation curve – which represents a trade-off between wages and employment rates– is likely a better specification for examining the wages effect on employment than a more traditional labor demand specification where the dependent variable would be the level of employment.

[31]This idea comes from Dahl (2002) who empirically tests a Roy (1951) model of self-selection of workers across states. He finds that while migration patterns of workers are partially motivated by comparative advantage, amenity differences across states also play a role in worker movements.

[32] The validity of the climate instruments rests on the assumption that the relationship between city climate and city-industry job creation and cost advantages (the Ωs and ϵs) is constant over time. In this case, the relationship is entirely captured in time-invariant city-specific effects that are differenced out of the estimating equation. This assumption may not be valid if the evolution of these advantages are related to long-term climate conditions.

4.2 Exploring Worker Heterogeneity

The model we developed in section 1 conceptually applies to workers of a single skill group. In section 1.4 we discussed how we address worker heterogeneity while at the same time focusing on the workforce as a whole. In this section, we assess this approach in two ways.

Our first approach is to evaluate the sensitivity of our results to including variables designed to capture changes in the skill distribution of a city's labor force. As discussed earlier, in a version of the model with perfectly substitutable bundles of efficiency units of work, the skill distribution of the labor force can affect job creation. We argued that we expect this effect to be second order in size but we wish to assess that expectation. The first variable we construct is the average efficiency units per person in a city. We calculate efficiency units for a person type defined by experience, education, gender and immigrant status as the average real wage over our entire period for people of that type, and then assign efficiency units to each person in a city based on their type.[33] Increases in this variable indicate that the workers in a city have become, on average, more productive and is designed to capture the effect that a firm opening a new vacancy has a higher probability of meeting a higher quality worker. We also construct an instrument designed to capture plausibly exogenous movements in this variable. To do this, we draw on literature that uses "enclaves" to predict immigration flows. In particular, we exploit the fact that immigrants from different sending countries have, on average, different educational attainment as well as other observable characteristics that influence a worker's efficiency weight.

We construct this instrument as follows: we assume immigrants from sending country h entering the country over a particular decade, denoted by M_{ht}, choose cities based to some extent on where previous waves of immigrants from the same sending country had settled. Denote the fraction of immigrants from sending country h living in city c at time t, as λ_{hct}. The predicted number of immigrants that will move to city c in year t, $\hat{M}_{ct}$, can be written as:

$$\hat{M}_{ct} = \sum_h \lambda_{hct-1} \cdot M_{ht}.$$

Similarly, we can construct the predicted number of workers in efficiency units that move to city c in year t as $\hat{M}^{EU}_{ct} = \sum_h \lambda_{hct-1} \cdot M_{ht} \cdot \omega_{ht-1}$, where ω_{ht-1} is the efficiency weight per worker from sending country h in the base year. To predict changes in the average efficiency units per worker, assuming no other changes in population, we construct the following instrument:

$$IV^{EU}_{ct} = \left(\frac{\hat{M}^{EU}_{ct}}{\hat{M}_{ct}} - EU_{ct-1} \right) \cdot \frac{\hat{M}_{ct}}{\hat{M}_{ct} + L_{ct-1}}, \tag{19}$$

where $\frac{\hat{M}^{EU}_{ct}}{\hat{M}_{ct}}$ is the predicted average efficiency units per worker arriving in city c in year t and EU_{ct-1} is the actual average efficiency units in city c in year $t-1$.[34]

[33] For additional details on how we calculate efficiency units, see Appendix C.

[34] The formulation of this instrument is similar to Doms and Lewis (2006), and in the same spirit as Card and DiNardo (2000) and Card (2001, 2009), among others.

The results that include changes in efficiency units per worker as an additional control are presented in Columns 1-4 of Table 5. In the first column, where all variables are treated as exogenous, the change in efficiency units per worker does not enter significantly. IV results are given in Columns 2-4, where we include IV_{ct}^{EU} in the instrument set. In line with previous literature, the immigration enclave instrument predicts changes in a city's skill composition quite well.[35] Nevertheless, in no specification is the skill variable significant, and, more importantly, it alters the magnitude of the estimated demand elasticities very little. To further probe sensitivity of our results to changes in city skill distribution, Columns 5-8 add alternative measures of a city's skill. In Columns 5 and 6, we include the change in the fraction of college graduates as an additional control. When treated exogenously, as in Column 5, it enters significantly but does not change the main conclusions regarding demand elasticities. In Column 6, we instrument the change in the proportion of college graduates with an enclave instrument similar to (19).[36] This results in imprecise estimates of the skill effect that are not significantly different from zero despite a strong first-stage. In the remaining Columns, 7 and 8, we include the change in the proportion of workers with education greater than high school. Again, inclusion of this variable does not alter the main conclusions of the paper. We conclude that changes in a city's skill distribution have, at best, only a second order effect on job creation and do not appreciably affect the estimated magnitude of our key parameters.

Our second approach (presented in Table 6) is to estimate our basic specification separately by education group. The education groups we consider are those with high school education or less and those with some post secondary or more.[37] When we perform this exercise, we are assuming that there are two completely segregated markets defined by education.[38] The dependent variable in Table 6 is the change in log city-industry employment rates for a particular education group. Similarly, wages and their instruments are constructed separately by education group.[39] However, we maintain the use of aggregate, city-wide employment rates (over all education groups) to capture the effects of limited entrepreneurial talent (which should operate at the city level) and vacancy contact rates.[40] Columns (1)-(4) pertain to the low-

[35] The first-stage t-ratio on IV_{ct}^{EU} in the change in efficiency units per worker equation is over 8.0 in all specifications.

[36] In this case, the instrument is constructed in exactly the same way as Doms and Lewis (2006).

[37] We have assessed the sensitivity of our results to finer breakdowns in education which typically resulted in very imprecise estimates. Finer skill definitions dramatically reduces the number of city-industry cells to work with, and results in sample size problems.

[38] Empirical evidence suggests that workers within our education classes are perfect substitutes, but that there is imperfect substitution of workers between the high- and low-education groups (Card, 2009). This latter type of substitution is ruled out in this framework.

[39] For example, $IV1$ and $IV2$ are constructed using city-industry shares and national wage premia that are estimated off of education specific samples.

[40] While it may be argued that vacancy contact rates for a particular education group would vary with education specific employment rates, using education specific employment rates made very little difference in practice. In addition, to the extent that one believes vacancy contact rates play little role in job creation compared to limited entrepreneurial ability, the aggregate employment rate specification would be preferred.

education group and columns (5)-(8) to the high-education group. The results for the low education group are very similar to those for the full sample. The results for the (smaller) post-secondary group are more erratic but tend to imply a similar sized wage elasticity.

4.3 Further Model Implications

To this point we have focused on how changes in wages, driven by changes in the bargaining power of workers, affect the employment rate in a city. The main mechanism by which a change in wages affects employment in our model is through inciting individuals to become entrepreneurs and create jobs. When wages are high, the profit to creating jobs is low and fewer individuals should choose to become entrepreneurs. In such a case, we should observe a low ratio of entrepreneurs-to-workers as only very efficient individuals should become entrepreneurs. In this section we want to explore whether observable entrepreneurial activity is consistent with this mechanism. In particular we will examine how changes in wages induce changes in the ratio of entrepreneurs to employees. Our measure for entrepreneurs will be the number of self-employed individuals in an industry-city cell. While self-employment is not a perfect measure of entrepreneurship, it is the most commonly used measure in the literature and we follow this tradition here. We exclude professionals from our measure of self-employed as these are represent mainly by physicians and lawyers who do not likely reflect the mechanism emphasized in the model. Including the professionals does not affect the main result we describe below.

Recall from section 1.1 that the number of entrepreneurs in industry i in city c – denoted S_{ict} – can be expressed as

$$S_{ict} = \left(\int_{\frac{K}{V^v_{ict}}}^{\infty} f(n)dn \right) \cdot L_{ct} \cdot \Omega_{ict}. \tag{20}$$

To get an expression for the ratio of entrepreneurs-to-employees, which we will refer to as entrepreneurial intensity, we divide equation (20) by E_c. The log-linear approximation for $\frac{S_{ict}}{E_{ct}}$ can then be expressed as

$$\Delta \frac{S_{ict}}{E_{ct}} = \rho_{it} + \rho_2 \Delta \ln V^v_{ict} + \rho_3 \Delta \ln \frac{E_{ct}}{L_{ct}} + \Delta \Omega_{ict}, \tag{21}$$

where $\rho_2 > 0$ and $\rho_3 < 0$. We can now use equation (7) to get the following expression:

$$\Delta \frac{S_{ict}}{E_{ct}} = (\rho_{it} + \rho_2 \tilde{\alpha}_{it}) + (\rho_2 \alpha_2) \Delta \ln w_{ict} + (\rho_3 + \rho_2 \alpha_3) \Delta \ln \frac{E_{ct}}{L_{ct}} + \rho_2 \alpha_4 \Delta \epsilon_{ict} + \Delta \Omega_{ict} \tag{22}$$

where $(\rho_2 \alpha_2)$ and $(\rho_3 + \rho_2 \alpha_3)$ are predicted to be negative. This equation will not be estimated consistently by OLS since we would expect the productivity changes in the error term to be positively correlated with wage changes. As the error term in (22) is of the same structure as in (7), the set of instruments we used to estimate the job creation equation (Z_1, Z_2 and Z_3) will provide consistent estimates of (22) under the same identifying assumptions used earlier.

Table 7 presents estimates of Equation (22). Column 1 contains estimates of the relationship obtained by OLS. The OLS estimated effect of wages on entrepreneurial intensity is close to zero, but as noted above, we expect this estimate to be biased. In Column 2 we use Z_3 and Z_1 to instrument the two regressors. In Columns 3 we use Z_2 and Z_1, and in Column 4 we Z_1, Z_2 and Z_3. Importantly, using any of these combinations of instruments, the estimated effect of wages on entrepreneurial intensity is significantly negative. Moreover, the coefficient is very stable across the three columns, with the over-id test reported in Column 4 showing that the over-identifying restriction is easy accepted by the data. As in our estimates of the job creation curve, we view the over-id test as a test of the validity of the model including our assumptions for the driving forces. We view it as quite telling that we find evidence in favor of the model even when we switch to a very different dependent variable. While the effect of wages on self-employment is quite precisely estimated, the estimated effect of market tightness is imprecisely estimated. The observed negative effect of wages on entrepreneurial activity provides some direct evidence in support the idea whereby higher wages reduce employment by reducing the incentive for individuals to become entrepreneurs and create jobs.

One of the key assumptions underlying our analysis is that entrepreneurs are proportional to the population, and that increasing the population in a city tends simply to lead to a replication of the employment structure in the city. This assumption was important for the results that employment determination should be stated in terms of employment rates instead of employment levels. This assumption also implies that entrepreneurial intensity should be unaffected by an exogenous increase in the population. In Columns 5 and 6 of Table 7 we explore this implication by adding labor force growth as an additional regressor to equation (22). In these three columns we instrument the wage and the employment rate using $IV1$, $IV2$ and $IV3$ as instruments. In Column 5 we do not instrument the labor force growth rate. In Column 6 we instrument labor force growth using the climate variables presented earlier as instruments. In both of the specifications, the effect of population growth is very small and not statistically significant at conventional levels. This supports the idea that entrepreneurs may be a scarce factor of production but that their supply appears to increase in proportion close to that of the population.

The combination of the result that entrepreneurs are proportional to the population and the result that population size does not help determine employment rates have important implications for addressing the puzzle set out at the begining of the paper: that is, why estimates using changes in policy parameters tend to imply inelastic labor demand curves while estimates based on migration shocks tend to imply very elastic curves. The policy shocks correspond to changes in labor costs and so, would be expected to trace out the slope of the job creation curve. In this sense, the fact that we obtain similar estimates from very different variation is encouraging. On the other hand, with a constant returns to scale matching function and entrepreneurs proportional to the population, an increase in the size of the population will not affect the equilibrium in employment rate - wage space since under these assumptions what matters is the tightness of the labor market not its size. Thus, when an immigration shock hits a local market we should see employment rise enough to restore the employment rate to its equilibrium level and the wage rate remain unchanged (after a period of adjustment). Plotted in employment level - wage space, the resulting relationship would be very flat and might be interpreted as revealing a very elastic labor demand curve. In the search and bargaining context, instead, this is simply a by-product of the main adjustments witnessed in employment rate – wage space. While much attention has been paid to David Card's result that the Mariel boat-lift did not alter wages in the Miami labor market, less has been made of his findings that the employment rate also returned to its previous level within a few years. We view this combination as fitting well with what we find.

5 Conclusion

In this paper, we extend and estimate an empirically tractable version of a search and bargaining model. Our goal is to highlight the implications of such a model for labor demand and to provide estimates of the responsiveness of employment outcomes to changes in average wages. In a search and bargaining model, labor demand is determined by points along the job creation curve, which is implicitly defined by a zero marginal profit condition for the creation of a vacancy. According to this condition, when wages rise the tightness of the labor market must fall in order to maintain the value of creating jobs. Therefore, the job creation curve defines a relationship between wages and employment rates rather than employment levels, as in more standard set-ups. We argue that this seemingly small difference has substantial implications. Chief among these is that shifts in labor supply and labor costs can produce very different labor market outcomes; that is, supply and wage shocks do not imply movements along the same demand curve in a search and bargaining model. This insight has the potential to explain some of the differences in labor demand elasticities that are presented in the literature.

In order to empirically evaluate the relevance of the search and bargaining framework, we extend an otherwise standard Pissarides (2000) model in several directions. First, our model includes multiple local labor markets linked through trade and labor mobility. There is substantial variation in employment levels and rates across cities, allowing for the identification of city-level job creation curves. Second, we extend the model to include multiple sectors, which is crucial for our identification strategy. Using insights developed in Beaudry, Green, and Sand (2012a), we use predicted industrial composition shifts to identify movements in average city wages induced by shifts in the outside options of workers. In a search and bargaining model, improvements in the outside employment opportunities of workers increases the bargained wage across all sectors within a locality. Finally, we follow Fonseca, Lopez-Garcia, and Pissarides (2001) in extending the standard model to include heterogeneous talent across the population in terms of abilities to create jobs in different industries. With this extension, the job creation curve may be less than perfectly elastic because of both search costs and the limited availability of job creators.

A key feature of this model is that it implies tight testable implications, as well as an over-identifying restriction. We use U.S. Census data from 1970-2007 to investigate whether city-level labor market outcomes conform to these restrictions. Working mainly at the industry-city level, our approach relies on comparing industry level changes in employment rates between cities with different changes in within-industry wages. We look at effects over periods of 10 years and, therefore, our estimates should be interpreted as representing long-run labor market outcomes. Importantly, we find that the model's main over-identifying restriction is easily passed in the data. While this observation does not necessarily prove that the theory is correct, we nevertheless believe that it provides compelling evidence in favor of the model. Based on this, we view estimates derived from this model as a reliable basis for assessing the impact of wage changes on employment outcomes.

The model implies two types of labor demand elasticities. The first corresponds to the effect of an increase in the cost of labor on a particular sector's employment rate while holding the overall tightness of the labor market in the city constant. Our estimate of this partial equilibrium elasticity is close to -1. We also examine how the local labor market as a whole reacts to a general increase in wages, allowing for interaction across sectors through the availability of workers and job creators. For this aggregate elasticity we obtain an estimate of approximately -0.3. We view the aggregate elasticity as the relevant concept for most policy discussions and see our estimate as indicating a reasonable but relatively inelastic relationship. For example, our estimates indicate that to obtain an increase in the employment rate of 1%, it is necessary to cut wage cost by 3%. Our model suggests that the somewhat low elasticity partly reflects a weak propensity for individuals to become more entrepreneurial and create more jobs as labor costs are reduced. This could arise either because the supply of new entrepreneurs is quite inelastic or because the creation of more jobs by existing entrepreneurs is limited by span of control problems. We hope to explore such mechanisms in more depth in future research.

Finally, the model provides a simple way to reconcile different estimates of the elasticity of labor demand obtained from policy shifts, such as minimum wage or payroll tax changes, versus those obtained from examining the response of wages to supply shifts. In a neo-classical framework both these approaches should identify the slope of the labor demand curve. However, the empirical literature generally comes to very different conclusion depending on which variation is used; with studies based on labor supply shifts suggesting an almost infinite elasticity of labor demand while those based on policy shifts suggest a much smaller elasticity. Within a search and bargaining model, these two sources of variation have quite different implications. On the one hand, a payroll tax should contribute to an increase in labor costs and, thus, can be used to identify the slope of the job creation curve. In this sense, it is encouraging that our estimated elasticity of the job creation curve broadly fits with the studies examining the effects of such policy interventions. On the other hand, an increase in the size of the labor force in the search and bargaining model does not trace out the equivalent of the labor demand curve. Instead, an increase in labor force size (perhaps due to immigration) should not change the equilibrium employment rate or wage, but should simply lead to an increase in the level of employment and an increase in the number of job creators. This is precisely the pattern we find when looking at cross-city outcomes.

References

ACEMOGLU, D. (1999): "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence," *American Economic Review*, 89(5), 1259–1278.

——— (2001): "Good Jobs versus Bad Jobs," *Journal of Labor Economics*, pp. 1–22.

BARTELSMAN, E. J., AND M. DOMS (2000): "Understanding productivity: lessons from longitudinal microdata," Finance and Economics Discussion Series 2000-19, Board of Governors of the Federal Reserve System (U.S.).

BARTIK, T. (1993): "Who Benefits from Local Job Growth: Migrants or the Original Residents?," *Regional Studies*, 27(4), 297–311.

BARTIK, T. J. (2006): "How Do the Effects of Local Growth on Employment Rates Vary With Initial Labor Market Conditions," Staff Working Papers 09-148, W.E. Upjohn Institute for Employment Research.

BEAUDRY, P., D. GREEN, AND B. SAND (2012a): "Does Industrial Composition Matter? An empirical evaluation based on Search and Bargaining Theory," *Econometrica*, (30), 1063–1104.

BEAUDRY, P., D. A. GREEN, AND B. M. SAND (2012b): "Spatial Equilibrium with Unemployment and Wage Bargaining: Theory and Estimation Wages," Working paper, University of British Columbia.

BERTOLA, G. (1992): "Labor Turnover Costs and Average Labor Demand," *Journal of Labor Economics*, 10(4), 389–411.

BLANCHARD, O. (1998): "Flows, Bargaining and Unemployment," Unpublished Manuscript, Department of Economics, M.I.T.

BLANCHARD, O. J., AND P. DIAMOND (1989): "The Beveridge Curve," *Brookings Papers on Economic Activity*, 20(1989-1), 1–76.

BLANCHARD, O. J., AND L. F. KATZ (1992): "Regional Evolutions," *Brookings Papers and Economic Activity*, pp. 1–75.

BLAU, F. D., AND L. M. KAHN (1999): "Institutions and laws in the labor market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 25, pp. 1399–1461. Elsevier.

BOUSTAN, L. P., P. V. FISHBACK, AND S. E. KANTOR (2007): "The Effect of Internal Migration on Local Labor Markets: American Cities During the Great Depression," NBER Working Papers 13276, National Bureau of Economic Research, Inc.

BROWN, J. N., AND O. ASHENFELTER (1986): "Testing the Efficiency of Employment Contracts," *Journal of Political Economy*, 94(3), S40–S87.

CARD, D. (1990): "The impact of the Mariel boatlift on the Miami labor market," *Industrial and Labor Relations Review*, 43(2), 245–257.

——— (2001): “Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration,” *Journal of Labor Economics*, 19(1), 22–64.

——— (2009): “Immigration and Inequality,” NBER Working Papers 14683, National Bureau of Economic Research, Inc.

CARD, D., AND J. DINARDO (2000): “Do Immigrant Inflows Lead to Native Outflows?,” *American Economic Review*, 90(2), 360–367.

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.

DOMS, M., AND E. LEWIS (2006): “Labor supply and personal computer adoption,” Working Papers 06-10, Federal Reserve Bank of Philadelphia.

FONSECA, R., P. LOPEZ-GARCIA, AND C. A. PISSARIDES (2001): “Entrepreneurship, start-up costs and employment,” *European Economic Review*, 45(4-6), 692 – 705.

HAMERMESH, D. (1993): *Labor Demand*. New Jersey:Princeton University Press.

JOVANOVIC, B. (1982): “Selection and the Evolution of Industry,” *Econometrica*, 50(3), 649–670.

KATZ, L. F., AND K. M. MURPHY (1992): “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *The Quarterly Journal of Economics*, 107(1), 35–78.

KRUEGER, A. B., AND J.-S. PISCHKE (1997): “Observations and Conjectures on the U.S. Employment Miracle,” NBER Working Papers 6146, National Bureau of Economic Research, Inc.

LEWIS, E. (2004): “How did the Miami labor market absorb the Mariel immigrants?,” Working Papers 04-3, Federal Reserve Bank of Philadelphia.

MACURDY, T. E., AND J. H. PENCAVEL (1986): “Testing between Competing Models of Wage and Employment Determination in Unionized Markets,” *Journal of Political Economy*, 94(3), S3–S39.

MANSKI, C. F. (1993): “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies*, 60(3), 531–42.

MELITZ, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, (6), 1695–1725.

MOFFITT, R. A. (2001): “Policy Interventions, Low-Level Equilibria And Social Interactions,” in *Social Dynamics*, pp. 45–82. MIT Press.

MORETTI, E. (2010): “Real Wage Inequality,” Unpublished manuscript, University of California at Berkeley, Department of Economics.

PARK, J. H. (1994): "Estimation of Sheepskin Effects and Returns to Schooling Using he Old and the New CPS Measures of Educational Attainment," Papers 338, Princeton, Department of Economics - Industrial Relations Sections.

PISSARIDES, C. (2000): *Equilibrium Unemployment Theory*. Cambridge, MA: MIT press, second edn.

ROBACK, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90(6), 1257–78.

ROGERSON, R., R. SHIMER, AND R. WRIGHT (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43(4), 959–988.

ROY, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, pp. 136–146.

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].

Table 1: First Stage Results

	(1)	(2)	(3)
$IV1$	3.370*		2.378*
	(0.651)		(0.635)
$IV2$		3.273*	2.894*
		(0.448)	(0.412)
$\sum_i \eta_{i,c,t-1} \cdot g_{it}$	0.0237	0.264*	0.0513
	(0.0728)	(0.0648)	(0.0759)
Observations	33993	33993	33993
R^2	0.485	0.493	0.497

NOTES: Standard errors, in parentheses, are clustered at the city-year level. (*) denotes significance at the 5% level. All models estimated on a sample of 152 U.S cities using Census and ACS data for 1970-2007. The dependent variable is the decadal change in log city-industry wages.

Table 2: Estimates of Job Creation Equation (9)

	OLS	IV		
	(1)	(2)	(3)	(4)
$\Delta \log w_{ict}$	0.125*	-1.022*	-0.930*	-0.953*
	(0.0158)	(0.279)	(0.231)	(0.221)
$\Delta \log ER_{ct}$	0.812*	-1.831*	-2.115*	-1.983*
	(0.0505)	(0.857)	(0.792)	(0.746)
Observations	33984	33984	33984	33984
R^2	0.506			
FR1		21.22		29.42
FR2			38.03	9.634
FEPC				
Over-id. p-val		.	.	0.676

NOTES: Standard errors are in parentheses. (*) denotes significance at the 5% level. All models estimated on a sample of 152 U.S cities using Census and ACS data for 1970-2007. Dependent variable is the decadal change in log city employment rate.

Table 3: Estimates of Aggregate Job Creation Equation (12)

	OLS	IV		
	(1)	(2)	(3)	(4)
$\Delta \log w_{ct}$	0.126*	-0.311*	-0.265*	-0.279*
	(0.0311)	(0.129)	(0.0896)	(0.0817)
$\sum_i \eta_{i,c,t-1} \cdot g_{it}$	0.0917*	0.219*	0.205*	0.210*
	(0.0350)	(0.0565)	(0.0440)	(0.0445)
N	608	608	608	608
R^2	0.497			
F-stat($\Delta \log w_{ct}$)		33.79	59.48	39.95
Over-id(p-val.)		.	.	0.728

NOTES: Standard errors are in parentheses. (*) denotes significance at the 5% level. All models estimated on a sample of 152 U.S cities using Census and ACS data for 1970-2007. Dependent variable is the decadal change in log city employment rate.

Table 4: Assessing the Impact of City Size

	OLS	IV				
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log w_{ict}$	0.138*	-1.008*	-0.912*	-0.933*	-0.773*	
	(0.0161)	(0.327)	(0.246)	(0.241)	(0.223)	
$\Delta \log \frac{E_{ct}}{L_{ct}}$	0.817*	-1.794*	-2.027*	-1.911*	-1.286*	
	(0.0480)	(0.824)	(0.833)	(0.765)	(0.620)	
$\Delta \log w_{ct}$						-0.270*
						(0.0883)
$\sum_i \eta_{i,c,t-1} \cdot g_{it}$						0.252*
						(0.0657)
$\Delta \log L_{ct}$	-0.101*	-0.00742	-0.0132	-0.0122	-0.115	-0.0373
	(0.0107)	(0.0425)	(0.0369)	(0.0369)	(0.0721)	(0.0374)
N	33993	33993	33993	33993	33993	608
R^2	0.508					
F-stat($\Delta \log w_{ict}$ or $\Delta \log w_{ct}$)		15.44	32.93	25.97	15.21	14.69
F-stat($\Delta \log \frac{E_{ict}}{L_{ct}}$)		8.577	12.00	9.443	5.096	.
F-stat($\Delta \log L_{ct}$)					24.00	12.34
Over-id(p-val.)		.	.	0.697	0.996	0.998

NOTES: Standard errors, in parentheses, are clustered at the city-year level. (*) denotes significance at the 5% level. All models estimated on a sample of 152 U.S cities using Census and ACS data for 1970-2007. Dependent variable is the decadal change in the log city-industry employment rate.

Table 5: Assessing the Impact of City Skill

	OLS	IV			OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta \log w_{ict}$	0.125*	-1.037*	-0.774*	-0.849*	-1.161*	-1.308*	-0.698*	-0.735*
	(0.0157)	(0.270)	(0.272)	(0.221)	(0.285)	(0.447)	(0.189)	(0.217)
$\Delta \log \frac{E_{ct}}{L_{ct}}$	0.801*	-2.092*	-2.999*	-2.329*	-2.355*	-2.435*	-1.842*	-1.876*
	(0.0529)	(0.926)	(1.229)	(0.891)	(0.937)	(1.193)	(0.700)	(0.708)
$\Delta \frac{\text{Efficiency Units}}{L_{ct}}$	0.0519	2.096	2.747	2.138				
	(0.0952)	(1.285)	(1.838)	(1.338)				
Δ BA or >					1.901*	3.366		
					(0.626)	(2.859)		
Δ SP or >							1.645*	1.406
							(0.419)	(0.816)
N	33993	33993	33993	33993	33993	33993	33993	33993
R^2	0.506							
F-stat($\Delta \log w_{ict}$)		14.22	25.41	22.08	28.93	22.15	27.70	22.03
F-stat($\Delta \log \frac{E_{ict}}{L_{ct}}$)		6.577	10.24	8.098	9.185	7.918	7.150	8.168
Fstat(Educ.Var.)		21.85	27.65	21.99	.	14.57	.	25.24
Over-id(p-val.)		.	.	0.292	0.377	0.268	0.108	0.197

Notes: Standard errors, in parentheses, are clustered at the city-year level. (*) denotes significance at the 5% level. All models estimated on a sample of 152 U.S cities using Census and ACS data for 1970-2007. Dependent variable is decadal changes in log city-industry employment rate.

Table 6: Results by Education Group

	HS or <				BA or >			
	OLS	IV			OLS	IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δw_{ict}	0.0614*	-0.881*	-0.928*	-0.911*	0.126*	1.659	-1.218*	-1.150*
	(0.0152)	(0.323)	(0.269)	(0.253)	(0.0191)	(8.120)	(0.217)	(0.210)
$\Delta \log \frac{E_{ct}}{L_{ct}}$	1.174*	-1.972	-1.830	-1.904*	0.220*	-10.36	-0.750	-0.928
	(0.0692)	(1.140)	(0.943)	(0.952)	(0.0670)	(28.44)	(0.629)	(0.655)
N	24375	24375	24375	24375	11651	11651	11651	11651
R^2	0.484				0.498			
F-stat($\Delta \log w_{ict}$)		20.89	25.84	24.75		11.52	28.39	21.35
F-stat($\Delta \log \frac{E_{ict}}{L_{ct}}$)		6.055	9.214	7.164		2.339	7.871	5.828
Over-id(p-val.)		.	.	0.874		.	.	0.147

Notes: Standard errors, in parentheses, are clustered at the city-year level. (*) denotes significance at the 5% level. All models estimated on a sample of 152 U.S cities using Census and ACS data for 1970-2007. Dependent variable is the decadal change in log city-industry employment rate.

Table 7: Estimates of Self Employment Equation (22)

	OLS	IV				
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log w_{ict}$	-0.020	-0.68*	-0.56*	-0.58*	-0.62*	-0.47
	(0.038)	(0.32)	(0.24)	(0.25)	(0.29)	(0.27)
$\Delta \log ER_{ct}$	0.0026	-0.49	-0.90	-0.71	-0.84	-0.46
	(0.15)	(0.85)	(0.93)	(0.82)	(0.90)	(0.67)
$\Delta \log$ Labor Force					0.024	-0.084
					(0.051)	(0.073)
Observations	21658	21658	21658	21658	21658	21658
R^2	0.30					
Instrument Set		Z_1, Z_2	Z_1, Z_3	Z_1, Z_2, Z_3	Z_1, Z_2, Z_3	Z_1, Z_2, Z_3,Climate
F-stat ($\Delta \log w_{ict}$)		16.10	33.36	25.81	22.32	11.31
F-stat($\Delta \log \frac{E_{ct}}{L_{ct}}$)		6.99	11.13	8.19	8.06	4.16
F-stat($\log L_{ct}$)						26.09
Over-id. p-val		.	.	0.55	0.54	0.29

NOTES: Standard errors, in parentheses, are clustered at the city-year level. (*) denotes significance at the 5% level. All models estimated on a sample of 152 U.S cities using Census and ACS data for 1970-2007. The dependent variable is the decadal change in the log Self Employment to Employment Ratio. All models inlcude an unrestricted set of year $\times$ industry dummies.

A Derivation of the Main Estimating Equation

In this appendix, we set out the details of the approximations used in deriving our main estimating equation. The first step is the log-linearization of equation (6), which we re-write here for convenience:

$$rV_{ic}^{v} = \alpha_{c1} \cdot h_i + \alpha_{c2} \cdot (p_i - w_{ic} + \epsilon_{ic}), \tag{A-1}$$

where $\alpha_{c1} = -\frac{(r+\delta)}{r+\delta+\phi_c}$ and $\alpha_{c2} = \frac{\phi_c}{r+\delta+\phi_c}$. Recall that ϕ_c can be expressed as a function of the city's employment rate, and we wish to make that relationship explicit.

We take a log-linear approximation of (6) around the point where cities have an identical employment rate and industrial composition (this arises when cities have $\epsilon_{ic} = 0$, and $\tilde{\Omega}_{ic} = \Omega_i$, where Ω_i is an industry level job creation cost that is common to all cities). This is given by:

$$\begin{aligned}\ln V_{ic}^{v} \approx \alpha_1 \cdot h_i + \alpha_2 \cdot (p_i - w_i) + \left[\frac{r+\delta}{(r+\delta+\phi)^2} \cdot (h_i + p_i - w_i)\right] \cdot \frac{\partial \phi}{\partial \frac{E}{L}} \cdot \frac{E}{L} \cdot \ln \frac{E_c}{L_c} - \ln \frac{E}{L} \\ - \alpha_2 \cdot (\ln w_{ic} - \ln(w_i) + \alpha_2 \cdot \epsilon_{ic}\end{aligned} \tag{A-2}$$

where, $\alpha_1 = -\frac{(r+\delta)}{r+\delta+\phi}$, $\alpha_2 = \frac{\phi}{r+\delta+\phi}$ and ϕ is the value for the probability that a vacancy meets a worker that is common to all cities at the base point.

Gathering the terms that are entirely at that national-industry level together and taking differences yields equation:

$$\Delta \ln V_{ic}^{v} = \tilde{\alpha}_i - \alpha_2 \Delta \ln w_{ic} + \left[\frac{r+\delta}{(r+\delta+\phi)^2} \cdot (h_i + p_i - w_i)\right] \cdot \frac{\partial \phi}{\partial \frac{E}{L}} \cdot \frac{E}{L} \cdot \Delta \ln \frac{E_c}{L_c} + \alpha_2 \epsilon_{ic} \tag{A-3}$$

The term multiplying $\Delta \ln \frac{E_c}{L_c}$ is called α_3 in equation (7) in the text and is negative because $\frac{\partial \phi}{\partial \frac{E}{L}} < 0$.

We also need to log-linearize the expression related to the entry condition for entrepreneurs, (5). This can be written as

$$\ln \frac{N_{ict}}{L_{ct}} = \theta_1 \ln V_{ict}^{v} + \theta_2 \tilde{\Omega}_{ict}, \tag{A-4}$$

where θ_1 is positive since as the value of vacancies increases the incentive to create jobs increases. As we wish to get an expression in terms employment instead of jobs, we use the equilibrium condition that $\delta \cdot E_{ic} = \phi_c \cdot [N_{ic} - E_{ic}] \quad \forall i$. Substituting for N_{ict} in (A-5) yields,

$$\ln \frac{E_{ict}}{L_{ct}} = \ln \left(\frac{\phi}{\delta + \phi}\right) + \theta_1 \ln V_{ict}^{v} + \theta_2 \tilde{\Omega}_{ict} \tag{A-5}$$

Then, differencing yields:

$$\Delta \ln \frac{E_{ict}}{L_{ct}} = \theta_1 \Delta \ln V_{ict}^{v} + \theta_2 \Delta \tilde{\Omega}_{ict} \tag{A-6}$$

Finally, by replacing $\Delta \ln V_{ict}^{v}$ with its expression given in (7), we can obtain an expression for job creation in an industry-city cell as a function of wages and overall labor market tightness:

$$\Delta \ln \frac{E_{ict}}{L_{ct}} = \theta_1 \tilde{\alpha}_{it} - \theta_1 \alpha_2 \Delta \ln w_{ict} + \theta_1 \alpha_3 \Delta \ln \frac{E_{ct}}{L_{ct}} + \theta_1 \alpha_4 \Delta \epsilon_{ict} + \theta_2 \tilde{\Omega}_{ict} \tag{A-7}$$

In the text this is written as equation (9), our main estimating equation:

$$\Delta \ln \frac{E_{ict}}{L_{ct}} = \varphi_{it} + \varphi_2 \Delta \ln w_{ict} + \varphi_3 \Delta \ln \frac{E_{ct}}{L_{ct}} + \zeta_{ict}. \tag{A-8}$$

B Consistency

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

$$\lim_{C,I \to \infty} \frac{1}{I} \frac{1}{C} \sum_i \sum_c Z_c \zeta_{ic} = 0. \tag{B-9}$$

We will handle the limiting arguments sequentially, allowing $C \to \infty$ first.[41] Recall that Z_{2c} is given by:

$$Z_{2c} = \sum_j \eta_{jc}(g_j^* - 1)(w_j - w_1), \tag{B-10}$$

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, (B-9) becomes:

$$\lim_{C,I \to \infty} \frac{1}{I} \frac{1}{C} \sum_j (w_j - w_1) \sum_c \eta_{jc}(g_j^* - 1) \sum_i \zeta_{ic}. \tag{B-11}$$

We can derive an equation for shares as:

$$\eta_{ic} = \frac{\tilde{\Omega}_{ic} \cdot F(V_{ic}^{v})}{\sum_i \tilde{\Omega}_{ic} \cdot F(V_{ic}^{v})} \tag{B-12}$$

Taking a linear approximation, again, around the point where $\Omega_{ic} = \epsilon_{ic} = 0$ yields the following expression:[42]

$$\eta_{ic} \approx \frac{1}{I} + \pi_1 \left(\epsilon_{ic} - \frac{1}{I} \sum_i \epsilon_{ic} \right) + \pi_2 \left(\Omega_{ic} - \frac{1}{I} \sum_i \Omega_{ic} \right) \tag{B-13}$$

The πs are positive coefficients obtained from linear approximation. We can decompose the ϵs and Ωs into absolute and comparative advantages, $\epsilon_{ic} = \hat{\epsilon}_c + v_{ic}^{\epsilon}$ and $\Omega_{ic} = \hat{\Omega}_c + v_{ic}^{\Omega}$, which gives:

$$\eta_{ic} = \frac{1}{I} + \pi_1 \cdot v_{ic}^{\epsilon} + \pi_2 \cdot v_{ic}^{\Omega} \tag{B-14}$$

[41] Throughout this appendix we omit the t subscript for simplicity.

[42] Recall that if $\Omega_{ic} = \epsilon_{ic} = 0$ then the industry shares are equal across cities. We further assume, for simplicity, that the shares are equal to $\frac{1}{I}$.

Similarly, substituting $\epsilon_{ic} = \hat{\epsilon}_c + v^{\epsilon}_{ic}$ and $\Omega_{ic} = \hat{\Omega}_c + v^{\Omega}_{ic}$ into the last term of (B-11), gives:

$$\sum_i \zeta_{ic} = \lambda_1 \cdot I \cdot \Delta\hat{\epsilon}_c + \lambda_2 \cdot I \cdot \Delta\hat{\Omega}_c, \tag{B-15}$$

which depends only on the increments of the absolute advantage components.

Substituting back into (B-11), yields:

$$\lim_{I\to\infty} \frac{1}{I}\sum_j (w_j - w_1)(g^*_j - 1) \lim_{C\to\infty} \frac{1}{C}\sum_c (\frac{1}{I} + \pi_1 \cdot v^{\epsilon}_{ic} + \pi_2 \cdot v^{\Omega}_{ic})(\lambda_1 \cdot I \cdot \Delta\hat{\epsilon}_c + \lambda_2 \cdot I \cdot \Delta\hat{\Omega}_c). \tag{B-16}$$

Thus, (B-16) equals zero provided that $E(\Delta\hat{\epsilon}_c) = E(\Delta\hat{\Omega}_c) = 0$, and that $\Delta\hat{\epsilon}_c$ and $\Delta\hat{\Omega}_c$ are independent of past values of the relative advantage components, v^{ϵ}_{ic} and v^{Ω}_c. In other words, general improvements in a city must be unrelated to past industry relative advantages.

Similarly, the relevant condition when using Z_{3c} is given by

$$\lim_{C\to\infty} \frac{1}{C}\sum_j \Delta(w_j - w_1)\sum_c \eta_{jc}\sum_i \zeta_{ic}, \tag{B-17}$$

and the same conditions ($\Delta\hat{\epsilon}_c$ to be independent of past values of v^{ϵ}_{ic} and of v^{Ω}_{ic}) ensure that this condition equals zero.

An important implication follows from this discussion. If the key identifying assumption underlying the IVs does not hold (i.e., changes in absolute advantage are not independent of past comparative advantages) then the two IVs will place different weight on the problematic correlation (between $\Delta\hat{\epsilon}_c$ and v^{ϵ}_{ic}). In particular, Z_2 use the weights ν_{it-1}), while Z_3 uses the weights $\Delta\nu_{it}$) and, therefore, estimates based on Z_2 and Z_3 should diverge.

Showing the proof is helpful for clarifying a potentially confusing point. Because national-level industrial growth rates (g_{it}) are aggregates of city-level growth and, therefore, of city-level productivity, one might be concerned that the presence of g^*_{it} in Z_2 creates a problematic correlation with the error term. But the expression in (B-16) makes it clear this is not the case. The reason is that (with the inclusion of a complete set of industry and time dummies) our identifying variation is cross-city, within-industry variation. That is, using Z_2, we are comparing employment rate changes in an industry between cities predicted to have more versus less growth in what we have called average rent. The changes in predicted rent are based partly on national-level growth rates, but those are the same for all cities and so do not introduce problematic variation. In terms of the proof, this can by seen from the fact that the g^*_{jt} and ν_{jt} terms are pulled outside the city-level summation and so are irrelevant for the consistency condition. This property will also apply to the other instruments we use which have a similar shift-share structure.[43]

[43] The type of instrument we use here is rather common in the labor literature. However we are not aware of any other formal discussion of when and under what conditions it is an appropriate instrument.

C Data (Not For Publication Appendix)

The Census data was obtained with extractions done using the IPUMS system (see Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander (2004)). The files were the 1980 5% State (A Sample), 1990 State, 2000 5% Census PUMS, and the 2007 American Community Survey. 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.[44] 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 courser 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 recoded variable `IND1950`, which recodes census industry codes to the 1950 definitions. This generates 144 consistent industries.[45] 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.[46] We are

[44]See http://www.census.gov/population/estimates/pastmetro.html for details.

[45]See http://usa.ipums.org/usa-action/variableDescription.do?mnemonic=IND1950 for details.

[46] The program used to convert 1990 codes to 1980 comparable codes is available at

also able to define more (231) consistent cities for that period.

We create a variable to proxy for the cost of housing in a city by using a measure of the rental rate of a two or three bedroom apartment in that city. To construct this variable, we use the Census variable for 'contract rent' and restrict it to the reported rent for two or three bedroom apartments in each of our MSAs. This is a similar procedure to that used recently by Moretti (2010). See that paper for a discussion on the appropriateness of this measure for local housing costs.

We construct efficiency units in the same manner as Katz and Murphy (1992). We begin by dividing the data, in each year, into demographic cells defined by potential experience (5 categories), education (4 categories), gender, race and immigrant status. We then calculate the average weekly wage and share of employment for each cell by year. We create a set of fixed-weights by averaging the employment shares for each demographic group across years, and construct an index of wages by year by taking the fixed-weighted average of wages in each year. We then divide cell-by-year wages by this index, and average across years for each demographic cell to obtain our measure of efficiency weights. When calculating efficiency units of workers, we use the efficiency weight multiplied by the sampling weight for each worker.

C.1 Enclave Instrument (Not for Publication Appendix)

The construction of the enclave instrument is similar to that of Doms and Lewis (2006) and uses their origin country groupings. The country of origin groups are (1) Mexico, (2) Central America, (3) South America, (4) Central Europe and Russia, (4) Caribbean, (5) China, (6) South East Asia, (7) India, (8) Canada, U.K., and Australia, (10) Africa, (11) Korea and Japan, (12) Pacific Islands, (13) Israel and NW Europe, (14) Middle East, (15) Central Asia, (16) Cuba, and (17) Souther Europe and can be identified from the IPUMS variable `bpl "Birthplace [general version]"` . To identify the inflows of immigrants, we use the IPUMS variable `yrimmig "Year of immigration"`.

C.2 Climate Instrument

The city level climate variables were extracted from "Sperling's Best Places to Live"

(http://www.bestplaces.net/docs/DataSource.aspx). Their data is compiled from the National Oceanic and Atmospheric Administration. The variables we use in this paper are the average daily high temperatures for July and January in degrees Fahrenheit. Alternative variables available from the same source are annual rainfall in inches and a comfort index. The comfort index is a variable created by "Sperling's Best Places to Live" that uses afternoon temperature in the summer and local humidity to create an index in which higher values reflect greater "comfort". We have

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 also compiled climate data from an alternative source to use as a robustness check. These data come from "CityRating.com's" historical weather data, and include variables on average annual temperature, number of extreme temperature days per year, humidity, and annual precipitation. Data from this source could only be collected for 106 cities, and, therefore, not included in this analysis.

D Worker Heterogeneity

Consider a simple search and bargaining model in which there are two types of workers: high- and low-skill. Let θ denote the fraction of high-skill workers and $\eta_H > 1$ and $\eta_L = 1$ denote productivity of high $\{H\}$ versus low $\{L\}$ skill workers.

The Bellman equations for firms can be written as:[47]

$$\rho V^v = -r + \phi \cdot \left(\theta \cdot V_H^f + (1-\theta) \cdot V_L^f - V^v\right) \tag{D-18}$$

$$\rho V_i^f = \eta_i \cdot p - w_i + \delta \cdot \left(V^v - V_i^f\right), \tag{D-19}$$

where p is the price of the output, and w_i is the wage paid for $i \in \{H, L\}$.

Combining (D-18) and (D-19), gives the value of a vacancy:

$$\rho V^v = \frac{-(\rho+\delta) \cdot r + \phi\left(\theta \cdot (\eta_H \cdot p - w_H) + (1-\theta) \cdot (p - w_L)\right)}{\rho + \delta + \phi}. \tag{D-20}$$

Workers of each skill type of the following Bellman equations:

$$\rho U_i^u = b + \psi \cdot (U_i^e - U_i^u) \tag{D-21}$$

$$\rho U_i^e = w_i + \delta \cdot (U_i^u - U_i^e), \tag{D-22}$$

for type $i \in \{\text{High}, \text{Low}\}$. Wages are set according to the rule $(1-\beta)(U_i^e - U_i^u) = \beta(V_i^f - V^v)$. This gives

$$w_i = (1-\beta) \cdot \rho U_i^u + \beta \cdot (\eta_i p - \rho V^v). \tag{D-23}$$

The Bellman equations for workers gives ρU_i^u as

$$\rho U_i^u = \frac{b \cdot (\rho + \delta) + \psi \cdot w_i}{\rho + \delta + \psi}. \tag{D-24}$$

Hence, wages can be written as:

$$w_i = \beta \Psi (\eta_i \cdot p - \rho V^v) + \frac{(1-\beta) \cdot \Psi \cdot b(\rho + \delta)}{\rho + \delta + \psi} \tag{D-25}$$

for each type where $\Psi = \frac{\rho+\delta+\psi}{\rho+\delta+\beta\psi}$. Then the wage differential between high- and low-skill workers is:

$$w_H - w_L = \beta \Psi p \cdot (\eta_H - 1) > 0. \tag{D-26}$$

[47] Here, we assume that firms meeting either high- or low-skill workers will form matches. This can be seen as a restriction that η_H cannot be too large.

We can also calculate the difference in the flow profits for a firm meeting a high-versus low-skill worker:

$$(\eta_H \cdot p - w_H) - (p - w_L) = p \cdot (\eta_H - 1) \cdot (1 - \beta\Psi) > 0 \tag{D-27}$$

Notice that we can write the high-skill wage in terms of the low-skill wage, which rationalizes the use of regression adjusted wages. Also, we can rearrange the job creation equation so that it depends on low-skill wage, plus θ times the difference in flow of profits. This reduces the value of a vacancy to a function of three variables $\rho V^v = G(\theta, \frac{E}{L}, w_L; \beta, \eta_i, \rho, \delta)$ where $\phi = \phi\left(\frac{E}{L}\right)$ and $\psi = \psi\left(\frac{E}{L}\right)$ are a function of the employment rate.

Setting $r = 0$, we can write

$$\rho V^v = \frac{\phi}{\rho + \delta + \phi} \cdot (\theta \cdot p \cdot (\eta_H - 1) \cdot (1 - \beta\Psi) + p - w_L). \tag{D-28}$$

The term interacting with the skill variable, θ, in equation (D-28) will be proportional to $(1 - \beta\Psi)$, which is equal to:

$$1 - \beta\Psi = \frac{(1 - \beta)(\rho + \delta)}{\rho + \delta + \beta\psi} \tag{D-29}$$

If $\rho+\delta$ is small relative to ψ (Blanchard, 1998), this term will also be small. Hence, we assume these skill effects only secondary and we can focus on $\rho V^v = \tilde{G}(\frac{E}{L}, w_L; \beta, \eta_i, \rho, \delta)$. This justifies using efficiency wages while not making any adjustments for $\frac{E}{L}$ in the baseline empirical work. We assess the sensitivity of this assumption in a robustness section by including measures of city skill, skill breakdowns, and efficiency units. None of our results seem to be sensitive to any of these alternative specifications.

E Worker Mobility (Not for Publication Appendix)

The purpose of this section is to extend our search and bargaining model to include worker mobility and to demonstrate that this extension does not change the main implications of our model. We consider two extensions of the model. In the first case, suppose unemployed workers have the option of occasionally switching cities. When this situation arises, with probability μ_1, workers choose to move to the city that maximizes their expected utility, $\max_{c'} U^u_{c'}$. To incorporate this extension, we can write workers' unemployment Bellman equation as:

$$\rho U^u_c = b + \tau_c + (1 - \mu_1)\psi_c \cdot \left(\sum_j \eta_{jc} U^e_{jc} - U^u_c\right) + \mu_1 \cdot (\max_{c'} U^u_{c'} - U^u_c), \tag{E-30}$$

where we have assumed that $\mu = 0$ to simplify exposition.[48] The outside options of workers are now changed and wage negotiation will take this into account. Importantly, the form of this change does not alter any of our results because $\max_{c'} U^u_{c'}$ does not depend on workers' initial location and, therefore, is captured in our empirical specifications by year-by-industry dummy variables. It should be noted that

[48]This is without loss and has the implication that unemployed flows depend only on city rather than industry-city.

in this case, however, the parameters we estimate will have a slightly different interpretation because they will depend on μ_1 (the ease of switching cities). Moreover, since in equilibrium workers will equalize expected utility across locations the term $\max_{c'} U^u_{c'} - U^u_c$ in equation E-30 will equal zero. Therefore, the equilibrium equations derived in the model section of this paper are not affected by this extension of worker mobility.

In the second extension, we include housing prices and local amenities, as in Roback (1982). We continue to allow workers to search across cities and choose the city that maximizes expected utility, as above. When doing so, workers take into account local housing costs and amenities. To incorporate this extension, assume that workers care about wages, the price of housing in a city, p^h_{ct}, 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, $w_{ic} - \vartheta p^h_{ct} + \Psi_c$. Accordingly, his or her flow utility when unemployed will be given by $b + \tau_c - \vartheta p^h_{ct} + \Psi_c$. The first thing to notice about this extension is that housing prices do not directly impact wage negotiation because housing costs are incurred in both the employed and unemployed state. In order for expected utilities to equalize across cities, housing prices must adjust. To capture this, we summarize the functioning of the housing market by assuming that housing prices can be expressed as a positive function of the population of a city and of amenities, given by:

$$p^h_{ct} = d_0 + d_1 \cdot L_{ct} + d_2 \cdot \Psi_c.$$

It is straight forward to derive an expression for housing prices, p^h_{ct}, that depends on local expected wages (and hence, on R_{ct}), amenities, and employment rates or labor market tightness. Housing prices will adjust such that a city with a favorable composition of jobs (due to the Ωs and ϵs) and higher amenities have benefits that are captured by local landowners. Wage differences across cities will not be equalized because in-migration will drive up local housing costs causing the movement of workers between cities to stop before nominal wage equalization occurs. We conclude this section by noting that the forces emphasized in our model have testable implications for labor mobility and housing prices, but do not alter the main conclusions of our baseline model.

F Selection Correction (Not for Publication Appendix)

The approach we use to address the issue of selection on unobservables of workers across cities follows Dahl (2002). Dahl argues that, under a sufficiency assumption, the selection-related error mean term in the wage equation for individual i can be expressed as a flexible function of the probability that a person born in i's state of birth actually chooses to live in city c in each Census year.[49] Dahl's approach is a

[49]This sufficiency assumption essentially says that knowing the probability of an individual's observed or "first-best" choice is all that is relevant for determining the selection effect, and that the probabilities

two-step procedure that first requires estimates of the probability that i made the observed choice and then adds functions of these estimates into the wage equation to proxy for the error mean term. Dahl also presents a flexible method of estimating the migration probabilities that groups individuals based on observable characteristics and uses mean migration flows as the probability estimates. We closely follow Dahl's procedure aside from several small changes to account for the fact that we use cities rather than states and to account for the location of foreign born workers.

Dahl's approach first groups observations based on whether they are ”stayers” or ”movers”. Dahl defines stayers as individuals that reside in their state of birth in the Census year. Since we use cities instead of states, we define stayers as those individuals that reside in a city that is at least partially located in individual's state of birth in a given Census year. Movers are defined as individuals that reside in a city that is not located in that individual's state of birth in a given Census year. We also retain foreign born workers, whereas Dahl drops them. For these workers, we essentially treat them as ”movers” and use their country of origin as their ”state of birth”.[50] Within the groups defined as stayers, movers, and immigrants, we additionally divide observations based on gender, education (4 groups), age (5 groups), black, and hispanic indicators. Movers are further divided by state of birth. For stayers, we further divide the cells based on family characteristics.[51] Immigrants are further divided into cells based on country of origin as described above.

As in Dahl (2002), we estimate the relevant migration probabilities using the proportion of people within cells, defined above, who made the same move or stayed in their birth state. For each group, we calculate the probability that an individual made the observed choice and for movers, we follow Dahl in also calculating the retention probability (i.e. the probability that individual i was born in a given state, and remained in a city situated at least partly in that state in general). For movers, the estimated probabilities that individuals are observed in city c in year t differ based on individuals' state of birth (and other observable characteristics). Thus, identification of the error mean term comes from the assumption that the state of birth does not affect the determination of individual wages, apart from through the selection term. For stayers, identification comes from differences in the probability of remaining in a city in ones birth state for individuals with different family circumstances. For immigrants, we assign the probability that an individual was observed in city c in a given Census year using the probabilities from immigrants with the same observable characteristics in the preceding Census year.[52] 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.

of choices that were not made do not matter in the determination of ones wage in the city where they actually locate.

[50]We use the same country of origin groups as for the enclave instrument.

[51]Specifically, we use single, married without children, and married with at least one child under the age of 5.

[52]For cities in the 1980 Census not observed in the 1970 Census, we use the 1980 probabilities.

Having estimated the observed choice or ”first-best” choice of stayers, movers, and immigrants and the retention probability for movers, we can then proceed to the second step in adjusting for selection bias. To do this, we add functions of these estimated probabilities into the first stage individual-level regressions used to calculate regression adjusted average city-industry wages. For movers, we add a quadratic of the probability that an observationally similar individual was born in a given state and was observed in a given city and a quadratic of the probability that an observationally similar individual stayed in their birth state. For stayers, we add a quadratic of the probability that an individual remained in their state of birth. For immigrants, we add a quadratic of the probability that an similar individual was observed in a given city in the preceding Census year. Dahl allows the coefficients on these functions to differ by state, whereas we assume that they are the same across all cities.