Title
Essays on the Economics of Space

Author
Bunten, Devin

Publication Date
2016

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

Essays on the Economics of Space

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Devin Michelle Bunten

2016

Abstract of the Dissertation

Essays on the Economics of Space

by

Devin Michelle Bunten

Doctor of Philosophy in Economics

University of California, Los Angeles, 2016

Professor Matthew Edwin Kahn, Chair

These essays contribute towards our understanding of the economics of space. This dissertation is composed of three chapters.

Chapter one—Is the rent too high? Aggregate implications of local land-use regulation: Highly productive U.S. cities are characterized by high housing prices, low housing stock growth, and restrictive land-use regulations (e.g., San Francisco). While new residents would benefit from housing stock growth due to higher incomes or shorter commutes, existing residents justify strict local land-use regulations on the grounds of congestion and other costs of further development. This paper assesses the welfare implications of these local regulations for income, congestion, and urban sprawl within a general equilibrium model with endogenous regulation. In the model, households choose from locations that vary exogenously by productivity and endogenously according to local externalities of congestion and sharing. Existing residents address these externalities by voting for regulations that limit local housing density. In equilibrium, these regulations bind and house prices compensate for differences across locations. Relative to the planner's optimum, the decentralized model generates spatial misallocation whereby high-productivity locations are settled at too-low densities. The model admits a straightforward calibration based on observed population density, expenditure shares on consumption and local services, and local incomes. Welfare and GDP would be 1.4% and

2.1% higher, respectively, under the planner's allocation. Abolishing zoning regulations entirely would increase GDP by 6%, but lower welfare by 5.9% due to greater congestion.

Chapter two—The impact of emerging climate risks on urban real estate price dynamics: In the typical asset market, an asset featuring uninsurable idiosyncratic risk must offer a higher rate of return to compensate risk-averse investors. A home offers a standard asset's risk and return opportunities, but it also bundles access to its city's amenities—and to its climate risks. As climate change research reveals the true nature of these risks, how does the equilibrium real estate pricing gradient change when households can sort into different cities? When the population is homogeneous, the real estate pricing gradient instantly reflects the "new news". With population heterogeneity, an event study research design will underestimate the valuation of climate risk for households in low-risk cities while overestimating the valuation of households in high-risk areas.

Chapter three—Entrepreneurship, Information, and Growth: We examine the contribution to economic growth of entrepreneurial marketplace information within a regional endogenous growth framework. Entrepreneurs are posited to provide an input to economic growth through the information revealed by their successes and failures. We empirically identify this information source with the regional variation in establishment births and deaths. To account for the potential endogeneity caused by forward-looking entrepreneurs, we utilize instruments based on historic mining activity. We find that the information spillover component of local establishment birth and death rates have significant positive effects on subsequent entrepreneurship and employment growth for U.S. counties and metropolitan areas. A version of this article was previously published in the Journal of Regional Science as Bunten et al. (2015).

The dissertation of Devin Michelle Bunten is approved.

Leah Michelle Boustan

Paavo Monkkonen

Pierre-Olivier Weill

Matthew Edwin Kahn, Committee Chair

University of California, Los Angeles

2016

To Ty–

the greatest of guys.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am immensely grateful for the many advisors, teachers, and coauthors I have been lucky enough to have found throughout my academic career. I am particularly indebted to the members of my dissertation committee, who I suspect have never all served on the same committee before: Matt Kahn, Leah Boustan, Pierre-Olivier Weill and Paavo Monkkonen, as well as the many others who have lent their guidance and support including Dora Costa, Walker Hanlon, and Lee Ohanian. My M.A. advisor and friend Stephan Weiler provided support throughout, and also lends his name to the third chapter. Thanks to Matt, Stephan, Eric Thompson, and Sammy Zahran, who are coauthors on the second (Matt) and third (Stephan, Eric, and Sammy) chapters. I benefited enormously from the comments of participants at the macro and economic history seminars at UCLA.

I am also eternally grateful to the fabulous friends and colleagues with whom I attended UCLA. In reverse alphabetical order, I am immensely grateful for the friendship, support, conversation, and insight I received from Rosanna Smart, Gaby Rubio, Sandra Rozo, Amanda Nguyễn, Owen Hearey, Federico Grinberg, Fer Giuliano, Sean Fahle, and many more classmates besides—especially those participants in the macro student group.

And of course, Tyanna Slobe listened to me ramble about zoning and cities for literally years until I had rambled myself into a dissertation, a degree, and a job. *Listening* here in the dynamic sense: you weren't just supportive but rather actively creative, and I quite literally could not have done this without you.

My research was also generously supported by Award Number T32AG033533 of the National Institute on Aging and Award Number T32HD007545 of the Eunice Kennedy Shriver National Institute Of Child Health & Human Development, through the California Center for Population Research at UCLA. The content does not necessarily represent the official views of any of the organizations that supported this research.

CURRICULUM VITAE

PROFESSIONAL EXPERIENCE

Board of Governors of the Federal Reserve System, Starting September 2016
Household and Business Spending|Division of Research and Statistics

RESEARCH AREAS

Urban Economics, Macroeconomics, Economic History, Public Economics

EDUCATION

University of California, Los Angeles, 2010–2016
M.A. in Economics, 2011

Colorado State University, 2008–2010
M.A. in Economics

University of Colorado at Denver, 2003–2005
B.A. in English/Film Studies, *magna cum laude*

AWARDS AND FELLOWSHIPS

Dissertation Year Fellowship, 2015–2016
Eugene V. Cota-Robles Fellow, 2010–2015
NIA/California Center for Population Research Trainee Fellowship, 2013–2014
NICHD/California Center for Population Research Trainee Fellowship, 2012–2013
UCLA Economics Department Outstanding Teaching Assistant, Principles of Microeconomics, 2012

RESEARCH

Is the Rent Too High? Aggregate Implications of Local Land-Use Regulations. Job Market Paper.

Entrepreneurship, Information, and Economic Growth, *Journal of Regional Science*. 2015. With Stephan Weiler, Eric Thompson, and Sammy Zahran.

Urbanization in the United States, 1800–2000. *Oxford Handbook of American Economic History*. Forthcoming. With Leah Boustan and Owen Hearey.

Optimal Real Estate Capital Durability and Localized Climate Change Disaster Risk. Working Paper with Matt Kahn.

The Golden Gated City: The Great Fire, Zoning and the San Francisco Housing Market. Work in progress with James Siodla.

CHAPTER 1

Is the Rent Too High? The Aggregate Implications of Local Land-Use Regulation

1.1 Introduction

Neighborhoods within productive, high-rent regions like the San Francisco Bay Area have very strict controls on housing development and very limited new housing construction. The Bay Area—home to Silicon Valley—is incredibly productive, with a GDP per capita of over $92,000, higher than all countries except tiny Luxembourg and oil-rich Norway and Qatar. Even for workers in non-tech sectors, wages in the region are higher than elsewhere in the US: median earnings for those with just a high school diploma are 12% higher in the Bay Area. Despite these high wages the region of 6.1 million people permitted just 20,046 new housing units in 2014, consistent with a growth rate of 0.8% annually. Of these, many were located in areas with few existing neighbors, either on former industrial land or previously undeveloped locations at the periphery.[1] The evidence suggests that locally-determined regulatory constraints are a substantial impediment to development (Glaeser and Gyourko, 2002, 2003; Glaeser et al., 2006). Existing residents justify these constraints by appealing to the costs of new development, including increased vehicle traffic and other types of congestion, and claim that they see few, if any, of the benefits

[1] In San Francisco, Mission Bay and the Dogpatch are exemplars of redevelopment on former industrial land (http://sf.curbed.com/archives/2014/05/29/mapping_25_projects_transforming_mission_bay_and_dogpatch.php).

from new development.[2] Instead, these benefits accrue to new potential residents, or to society at large.

While residents may pass local regulations to deal with local externalities of congestion, in aggregate the implications stretch beyond local borders: regions with highly-regulated municipalities experience less-elastic housing supply (Glaeser et al., 2006; Saiz, 2010). Land-use regulations in high-productivity cities constrain the expected supply response to sustained high rents, and these regulations are the result of local endogenous political processes.[3] Understanding the welfare effects of zoning regulations involves quantifying both the aggregate housing supply costs as well as the benefits that lead households to pass zoning regulations.

This paper builds a spatial equilibrium model of housing and endogenous regulation to assess the welfare implications of locally-determined zoning. To motivate households' preference for zoning, I introduce local externalities of agglomeration (e.g., residents can share fixed-cost infrastructure or a diversity of restaurants) and congestion (e.g., traffic or limited on-street parking) into a standard Rosen/Roback-style model of heterogeneous locations. In line with the externalities under study, I identify these locations with *neighborhoods*.[4] Existing residents feel the effects of these externalities from new development, but do not see the benefits of higher local output. They thus use local political processes to address these externalities by establishing *zoning laws* that limit development. The trade-off between agglomeration (i.e., sharing of fixed costs and services) and congestion ensures that households prefer a positive but limited number of co-residents. I use the model to identify the wedges between the zoning equilibrium and the planner's optimal

[2]For example, the group Livable Boulder supports a ballot initiative that they argue will "ensure that City levels of service are not diminished by new development", due to concerns about "huge buildings, blocked views of the mountains, more congestion, proposals to change the unique character of many Boulder neighborhoods" (http://livableboulder.org/).

[3]That regulations are determined locally—at the level of a neighborhood, city district, or municipality—is argued by Hills and Schleicher (2011, 2014); Schleicher (2013); and Monkkonen and Quigley (2008).

[4]Neighborhood productivity differences within and across cities arise from differential access to employment opportunities, transportation networks, or productive amenities.

allocation of households to locations. Compared to the utilitarian planner's allocation, high-productivity locations are developed less intensively, low-productivity locations are developed more intensively, and too many (low-productivity) locations are opened in a decentralized equilibrium.

The key model components—fixed costs, congestion, and endogenous regulation at the local level–are grounded in empirical evidence. The fixed costs of infrastructure and local services can explain the sharp edge of development visible in metropolitan areas, in contrast with the predictions of the standard monocentric city model of a smooth gradient approaching zero density at the fringe.[5] Local congestion resulting from fixed quantities of street parking, limited space for children's play, and concerns about sunshine, shadows, and wind permeate anti-development discourse.[6] Finally, zoning laws in most cities tend to be the result of neighborhood- or municipality-level political processes (Hills and Schleicher, 2014; Monkkonen and Quigley, 2008), rather than broad metropolitan outcomes.[7]

I model a large set of locations analogous to the neighborhoods or municipalities of the national economy. Opening a location to settlement necessitates payment of a fixed cost (e.g., of extending a sewer line) that can be shared more broadly with more residents. Balancing this incentive to share fixed costs, households in a location receive disutility from the addition of further co-residents. These forces of agglomeration and congestion, respectively, constitute the *endogenous amenities* provided by a location. In light of these forces, local residents choose *zoning laws* to restrict the maximum number of housing units that can be built by a competitive construction sector. Distributions of house prices, zoning laws, and location choices are jointly determined in equilibrium. Households are *ex*

[5] For a graphic example, see Figure A.2 in the appendix.

[6] A recent example from a zoning debate in Santa Monica is http://smdp.com/santa-monica-beach-town-dingbat-city/147819.

[7] Relatedly, Aura and Davidoff (2008) argue that housing and land demand elasticities imply that any particular municipality would be unable to increase supply sufficiently to move down housing demand curves to lower prices. The converse of this finding is that municipal (or neighborhood) governments have limited ability to act as monopolists and raise prices by simply restricting supply.

ante homogeneous and fully mobile, and so house prices adjust to ensure that households are indifferent between all occupied locations. In particular, productive locations have binding zoning restrictions that cause house prices to become significantly higher than the marginal cost of construction, an empirical finding documented in Glaeser and Gyourko (2002, 2003); Glaeser et al. (2005b,a); Cheshire and Hilber (2008) and Koster et al. (2012).

I next analyze the welfare implications of local zoning laws by comparing the equilibrium allocation with the optimal allocation chosen by a constrained planner.[8] When choosing the intensity of development within a location, the planner considers the local forces of agglomeration and congestion as well as the value of placing the marginal household in a location of higher or lower productivity. When passing zoning laws, households only consider the local amenity forces and thus restrict development too heavily in productive locations. This underdevelopment induces *sprawl*: the opening of too many unproductive locations, consistent with e.g. Fischel (1999). I further show that the planner's allocation can be implemented through a zoning reform where residents choose all zoning laws at the national level.

The model enables a straightforward calibration. I calibrate the utility function to the consumption share on non-local goods, excluding housing and local services. For the local sharing parameter, I calibrate to the share of spending on local services. The congestion cost curvature is chosen to match the average population density of urban census tracts.[9] For location productivity, I use Census tract income data, adjusted for average household observable characteristics (Hsieh and Moretti, 2015). I calibrate construction-sector productivity to match the ratio of price to construction costs in high-price locations like New York and San Francisco where the effects of zoning laws have been well-studied (cf. Glaeser et al., 2005b).

[8]The planner is *constrained* in that they maximize welfare subject to a spatial equilibrium constraint: identical households must receive identical utility, regardless of location.

[9]While the parameter governs the shape of the congestion cost function, in equilibrium it has a first-order effect on the optimal zoning laws and thus on average density. The approach is similar to the tradition of using wage shares to calibrate labor and capital exponents in a production function: in equilibrium, these shape parameters have first-order effects on levels.

The model provides a unified framework for addressing the full array of costs and benefits to reforming local zoning policy by implementing the optimal allocation. Avent (2011) and Hsieh and Moretti (2015) argue that these local growth patterns have been quite costly for national growth, while incumbent residents argue that new development would itself be costly. My main exercise is to quantify these costs while accounting for the changes to endogenous amenities resulting from the planner's allocation. The planner's allocation raises welfare and GDP by 1.4% and 2.1%, respectively. The more intensive development of productive locations necessary to increase GDP would also increase the congestion disamenity received by the residents of these regions. The median resident experiences an increase in density of 3.6%; increases as great of 10–15% are experienced by the most productive locations. More intensive development reduces the total number of locations opened to development by 3%. The planner's allocation thus features less *sprawl*, with less fixed-cost expenditure and commensurately higher consumption.

A second exercise offers a quantitative assessment of an alternative policy: zoning abolition, wherein developers are free to build housing without constraint.[10] This policy results in over-building in high-productivity areas as construction firms do not internalize congestion effects. Model GDP increases by 6%, but increased congestion causes welfare to decline by 5.9%. The extra productivity comes arises because productive regions like San Francisco see population inflows of 50% or more. The policy implication is that zoning should be relaxed in productive locations, but not abolished.

Treating zoning as endogenous enables the inference of welfare losses from zoning in the absence of detailed data on the wedges between housing prices and marginal costs of construction at the location level. Were such data available, the wedges could be treated as exogenous and the welfare costs of zoning would be the costs of moving the wedges from their observed to their optimal level. In principle, observed density could be used to proxy for underlying regulations. The relationship between productivity and zoned

[10] This is the option implicitly studied by Hsieh and Moretti, who find that output would increase by 13.5%.

density, however, is confounded in the data by factors like exogenous amenities—which make land more expensive and increase density—and long-lived investments, both in building stock and in infrastructure such as subway networks. Instead of attempting to account for these myriad forces, I model zoning as an equilibrium object, infer the unobserved wedge from the model, and calculate the loss in welfare accordingly. The key qualitative prediction of this model is that, in high-productivity areas, density does not rise fast enough relative to productivity.

Endogenous zoning has a further benefits: it offers predictions on the likely outcomes of different housing policy and zoning reform. First, policies which ignore the endogeneity of zoning may be unable to meet their stated goals. The federal government provides large housing subsidies throughout the income distribution, and increased may appear to offer an outlet for reducing the burden of high costs in expensive cities. Within the specified model, a housing subsidy would increase household willingness to pay in expensive areas. However, endogenous zoning driven by existing residents will not respond to these subsidies by creating more supply, and so the policy will not have the intended effect of making housing more affordable, nor will it stimulate additional supply in zoning-constrained locations.

Second, the model provides guidance for the outcomes of direct zoning reforms. Can a single location allow more development—increasing housing supply—and hope to bring down prices? Consider an infinitesimal neighborhood in the model that relaxes its zoning restriction and allows more development. The outside option—the value of locating elsewhere—will be unaffected, as a single location is insufficient to move the price gradient. Instead, house prices in the neighborhood will decline only to the extent that additional development makes congestion worse. This echoes the complaints of homeowners who fear new development will hurt their house values.[11]

[11]For example, This claim is offered by an advocate for a moratorium on development in the Mission District of San Francisco here: https://medium.com/@danancona/putting-market-fundamentalism-on-hold-432ecf1aab3c.

Finally, could a successful *upzoning*, where a set of productive locations increase the zoned capacity or even abolish zoning restrictions, harm landowners in less productive locales? Suppose that an entire productive region—perhaps a city, or a state—chooses to allow more intensive development. Many new households will flow in, abandoning the less productive locations they once occupied. For the least-productive locations, the initial population outflows will make the fixed costs too burdensome, and all households will choose to leave. If profits from development of a location are tied to that location, perhaps via homeownership, then residents of less productive locations may lose out from this policy reform.

1.1.1 Related Literature

This paper draws on work that has established a strong relationship between zoning and the elasticity of housing supply (Fischel, 1999; Mayer and Somerville, 2000; Glaeser and Gyourko, 2002; Glaeser et al., 2005b,a; Saiz, 2010; Koster et al., 2012; Turner et al., 2014).[12] This literature has shown that zoning laws restrict the supply response to high house prices, and that cities that do zone restrictively have systematically higher housing prices. Following this literature, zoning laws in this paper shape local amenities (cf. Turner et al., 2014) and regional housing supply (cf. Fischel, 1999; Mayer and Somerville, 2000).

Several authors in diverse contexts have modeled zoning as an endogenous outcome of political processes (Hamilton, 1975; Fischel, 1987; Monkkonen and Quigley, 2008; Fischel, 2009; Hilber and Robert-Nicoud, 2013; Ortalo-Magné and Prat, 2014; Hills and Schleicher, 2011; Schleicher, 2013; Hills and Schleicher, 2014; Fischel, 2015). I implement the findings of (e.g.) Fischel (2009) and Hills and Schleicher (2011) that zoning laws are actively determined by highly-engaged utility-maximizing households at the lo-

[12] Gyourko and Molloy (2014) present an in-depth review of the economics of housing regulation.

cal level—municipalities, neighborhoods, or even direct neighbors.[13] In some contexts, zoning enables households to implement the planner's allocation, usually by resolving problems of free-riding on local expenditure. I complement this literature by embedding a set of heterogeneous locations into a general equilibrium model, so that zoning is locally optimal but has aggregate external costs. Incorporating the endogeneity of local zoning enables a welfare calculation in the absence of detailed local data and also reflects the consensus of the literature (cf. Fischel, 2009, Hills and Schleicher, 2011, and Hills and Schleicher, 2014).

My work builds on Hsieh and Moretti (2015), who also link strict zoning and changes in aggregate output. They study the productivity effects of wage dispersion across metropolitan areas in a Rosen-Roback model, and attribute the increase in this dispersion (and resulting loss in output) to a decrease in the elasticity of housing supply. From this basis, I introduce local externalities and calculate welfare losses from zoning relative to the planner's optimum, in addition to output losses from spatial misallocation. My paper differs in terms of the economic interpretation of zoning: here, it is an endogenous limit on neighborhood development instead of a change in the elasticity of metropolitan housing supply. As noted, this endogeneity serves two purposes: it enables me to infer the costs of welfare in the absence of location-specific data on marginal costs and housing prices and it allows me to study the likely local outcomes of different policy reforms.

This paper is related to others that study the conditions under which zoning laws and other local goverment interventions may be theoretically optimal (Stull, 1974; Hamilton, 1975; Henderson, 1991; Hochman et al., 1995; Rossi-Hansberg, 2004; Calabrese et al., 2007; Chen and Lai, 2008; Allen et al., 2015). These papers focus on a variety of externalities to motivate zoning, including nuisance, production, and free-riding interactions that I abstract from. Of course, the optimal zoning scheme depends upon the nature of

[13] Lawsuits seeking to halt development in the name of various ills are increasingly common (Ganong and Shoag, 2013).

the externalities under consideration. Similarly to some of these papers, I characterize the zoning regime consistent with maximizing welfare given the externalities studied. This paper also addresses the optimal distribution of population between communities who share local public goods (Flatters et al., 1974; Arnott and Stiglitz, 1979; Hochman et al., 1995). Like Hochman et al. (1995), I find that jurisdictions that are well-positioned to deal with one externality may not be optimal in the presence of other inter-location effects.

More broadly, this paper relates to others that study the spatial determinants of aggregate productivity and welfare (Albouy, 2012; Desmet and Rossi-Hansberg, 2013b; Allen and Arkolakis, 2014; Behrens et al., 2014; Eeckhout and Guner, 2015; Morales et al., 2015; Hsieh and Moretti, 2015). I also account for the joint spatial distributions of household location choices, incomes, and housing prices while incorporating the effects of zoning regulations. As such, it relates to Glaeser et al. (2005a), Van Nieuwerburgh and Weill (2010), Gyourko et al. (2013), and Diamond (2016), among others.

The rest of the paper is organized as follows. Section 2 presents the model and the analytic results. Section 3 calibrates a quantitative version of the model to match features of the data. Section 4 quantitatively analyzes the welfare costs of local zoning and the no-zoning counterfactual relative to the planner's optimal allocation. Section 5 concludes.

1.2 Model

This section describes the model. I begin by outlining the environment. I then study the determination of equilibrium in three steps. First, I study the *local equilibrium* that results from the actions of households and firms, who take as given the local zoning rules as well as the endogenous outside option and level of profit. Second, I study the *local zoning choice* of initial residents. The zoning law restricts the level of housing production available to construction firms. Third, I study the economy-wide *general spatial equilibrium* in which the aggregate variables are determined: the endogenous outside-option level of utility and aggregate profits of construction firms.

Environment The spatial environment consists of a mass of locations of measure 1 with productivity indexed by $x \in [0,1]$. These locations represent the set of potential neighborhoods in the entire economy. The neighborhood with $x = 1$ is the most productive in the country, e.g. downtown San Francisco. Locations with x less than but close to 1 include less-productive locations in the same area—for instance, Oakland, a train ride from downtown San Francisco. A similar x could also correspond to the best locations in less-productive regions, like downtown Denver. Locations with x close to zero would be at the far fringes of metro areas: either a short commute to an unproductive job or a very long commute to a better job.

Production takes place at each location according to a linear production function. Productivity is given by a continuous function $y(x)$ with $y(0) = \underline{y}$, $y'(x) > 0$, and $y(1) = \bar{y}$. Each location begins empty and available for settlement by a measure 1 of households, whose preferences are described below. Opening a location to urban settlement requires a fixed cost F. The set of locations that are open in equilibrium is an endogenous outcome.[14]

I now turn to the description of agents and their maximization problems.

[14]As a technical matter, it is possible that fixed and construction costs are sufficiently large that there is no feasible distribution of households under which total output is sufficient to cover these costs. In that case the economy can be thought of as *agricultural*: no fixed costs are paid, and each location is occupied by a solitary household who produces $y_A < \underline{y}$. In the urban equilibrium, fewer locations will be opened. A final case is that the endogenous outside option of the urban economy is equal to the utility earned from agricultural, in which case some locations would remain agricultural. Restricting y_A to be sufficiently small effectively rules out this equilibrium.

1.2.1 Local Equilibrium

1.2.1.1 Household and Construction Firm Problems

The economy consists of three types of agents. First, there is a measure 1 of households who choose location and consumption to maximize utility, taking as given location-specific rents, population, and the endogenous outside option (i.e., the maximum attainable utility from making the optimal choice of location).

Second, there is a representative construction firm that chooses the housing production level to maximize at each location profits, taking as given rents and the local zoning law. I represent a zoning law as a maximum allowable number of housing units.[15]

Third, some locations begin with a measure 0 of *initial residents.* The initial residents choose local zoning laws and consumption to maximize utility, while taking as given a pre-fixed initial rent.[16] Initial residents have rational expectations about the equilibrium choices that households and firms will make given the zoning law. In locations with no initial residents, construction firms maximize profits without constraint.

I proceed by defining the problems of the household and the construction firm and then the neighborhood equilibrium given zoning laws. Then I define the problem of the initial residents and their equilibrium choices of zoning law.

Household Problem A household in location x has preferences summarized by the utility function

$$u(c(x)) - v(n(x)).$$

[15]As each household consumes a single housing unit, the choice of zoning law restricts both the number of housing units and the number of households.

[16]The initial rent can be thought of as the fixed mortgage payment of an existing homeowner, which does not fluctuate based on local rental conditions. It can also be though of directly as rent control, where the price cannot be raised on existing tenants. Rent control is a common policy tool in expensive rental markets like San Francisco.

The function $u(\cdot)$ represents preferences over consumption, $c(x)$. The function $v(\cdot)$ represents a congestion externality that depends on the neighborhood population, $n(x)$. Utility from consumption is twice continuously differentiable and concave: $u'(c) > 0$ and $u''(c) \leq 0$. The congestion externality $v(\cdot)$ is bounded below, increasing, twice continuously differentiable, and convex: $v'(n) > 0$ and $v''(n) \geq 0$. Without loss of generality, I assume that $v(0) = 0$.

The household faces budget constraint

$$c(x) = y(x) + \Pi - \frac{F}{n(x)} - p(x),$$

where $y(x)$ and $p(x)$ represent the location-specific income and house price. Consumption is given $c(x)$. Households own a diversified portfolio of construction firms in all locations, and Π is the location-independent transfer of construction sector profits from the national economy. The fixed cost F is shared equally among all residents.

Households take as given local prices $p(x)$, total profits Π, and their endogenous outside option $\bar{u}$. Households are freely mobile and choose location x and consumption $c(x)$ to maximize utility subject to the budget constraint.

Construction Firm Problem Construction firms choose the development intensity $n^f(x)$ to maximize profits at each location subject to a local zoning constraint $\bar{n}(x)$. There is no pre-existing housing stock in any location. Taking the price $p(x)$ as given, the firm problem in location x is

$$\max_{n^f(x)} \Pi(x) = p(x) n^f(x) - \left(\frac{n^f(x)}{Z} \right)^2$$

subject to zoning constraint

$$n^f(x) \leq \bar{n}(x) .$$

It is natural to think that construction costs within a location are convex due to the costs associated with building taller and denser structures. Construction firms earn positive profits from locations with positive construction. Profits are aggregated across all locations and redistributed equally to all households.[17]

1.2.1.2 Local Spatial Equilibrium

Consider a fixed $\bar{n}(x)$, a fixed $\bar{u}$, and a fixed Π. I define a *local spatial equilibrium* for location x to be a house price $p^*(x)$ and local population $n^*(x)$ such that households and firms solve their respective problems with households consuming their budget, the local housing market clears, and spatial equilibrium must hold. That is:

$$u\left(y(x) + \Pi - \frac{F}{n^*(x)} - p^*(x)\right) - v\left(n^*(x)\right) \leq \bar{u}, \tag{1.1}$$

with equality if $n^*(x) > 0$. Housing market clearing implies $n^f(x) = n^*(x)$. Using this, the firm's problem yields a complementary slackness condition which states that either the zoning constraint binds or price is equal to the marginal cost of construction:

$$\left(p^*(x) - 2n^*(x)/Z^2\right)\left[\bar{n}(x) - n^*(x)\right] = 0. \tag{1.2}$$

The firm must also abide by the zoning constraint:

$$n^*(x) \leq \bar{n}(x). \tag{1.3}$$

A *local spatial equilibrium* in location x is a pair $\{n^*(x), p^*(x)\}$ satisfying Equations (1.1)), (1.2)), and (1.3).

[17] The cost function is equivalent to a housing production function that is Cobb-Douglas in building materials and land, with a coefficient of 1/2 on each term. Under this alternative production function, the profit redistribution outlined here corresponds to an assumption that each household owns a diversified portfolio of land. This share is within the range of estimates provided by Albouy and Ehrlich (2012).

Housing Equilibrium Selection The existence of a fixed cost creates complementarities in household location decisions: if some households go to a location, the fixed cost can be shared more broadly and the location becomes more attractive to other households. This complementarity implies that there can be multiple equilibria. Finally, the congestion externality implies that eventually an additional household will lower the utility of existing households as the congestion costs outweigh the sharing benefits. I show that there are up to three pairs of equilibrium population $n^*(x)$ and price $p^*(x)$ that may satisfy the definition of an equilibrium. Figure 1.1 characterizes graphically the set of potential equilibria. The *optimal location* curve, labeled $OL(n,x)$, shows the set of (n,p) points consistent with the household's maximizing choices, and is defined as follows:

$$OL\left(n,x\right)=\begin{cases} y(x)+\Pi-\frac{F}{n}-u^{-1}\left(v\left(n\right)+\bar{u}\right) & \text{if } n>0 \\ [0,\infty] & \text{if } n=0. \end{cases} \tag{1.4}$$

The curve labeled $MC\left(n\right)$ corresponds to the marginal cost of construction and is independent of location:

$$MC\left(n\right)=2\;n/Z^{2}. \tag{1.5}$$

Proposition 1. *The portion of the* $OL\left(n,x\right)$ *curve with* $n(x)>0$ *is strictly concave, hump-shaped, and intersects the* $MC\left(n\right)$ *curve twice, not at all, or once at a point of tangency.*

Proof. See appendix. □

If the zoning constraint lies to the right of $n_H(x)$—the upper intersection of $MC(n)$ and $OL(n,x)$—then both $n_H(x)$ and $n_L(x)$ are consistent with equilibrium. If the zoning constraint lies within the interval $(n_L(x),n_H(x))$ then both $n_L(x)$—the lower intersection—and $\bar{n}(x)$ are consistent with equilibrium. If the zoning constraint lies to the left of $n_L(x)$, then the origin is the sole equilibrium. There is always an equilibrium

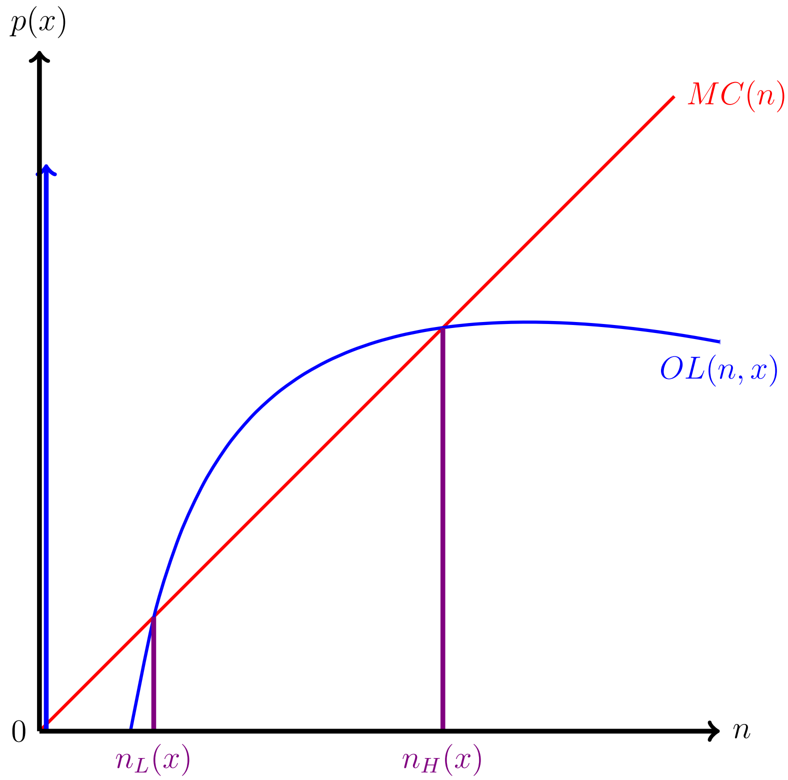

Figure 1.1: Marginal cost ($MC(n)$) and optimal location ($OL(n, x)$) curves. Note that the $OL(n, x)$ curve includes the entire vertical axis above 0. The set of points potentially consistent with equilibrium, for which $OL(n, x) > MC(n)$, is the closed interval $(n_L(x), n_H(x))$ and the point 0.

at the origin: households and firms may expect a $p^*(x) = 0$ and $n^*(x) = 0$ and see these expectations fulfilled.

If there are multiple equilibria, I select the one with the largest population.[18] This equilibrium selection has two virtues. First, it abstracts from coordination failures where both households and firms expect the location to be empty, and so it remains empty. Second, it is *stable* in the sense of economic geography models, i.e. a small deviation of population does not induce population movements that lead the location to a different

[18]Note that this supremum is taken over a closed and bounded set. By continuity of the utility and construction cost functions, the supremum of the set of points consistent with equilibrium is, itself, a member of this set.

equilibrium (cf. Krugman, 1991). This rules out $n_L(x)$ as an equilibrium population as it is would be unstable.[19] Concretely, both households and construction firms expect that all locations will be settled at the highest-population equilibrium.

The equilibrium population $n^*(x)$ and price $p^*(x)$ consistent with this selection criterion for a given zoning law $\bar{n}(x)$ are

$$\{n^*(x), p^*(x)\} = \begin{cases} \{0, 0\} & \text{if } MC\,(n) > OL\,(n, x)\,\forall\; n \in (0, \bar{n}(x)] \\ \{\bar{n}(x), OL\,(n, x)\} & \text{if } \bar{n}(x) \in [n_L(x), n_H(x)] \\ \{n_H(x), MC\,(n)\} & \text{if } \bar{n}(x) > n_H(x). \end{cases} \tag{1.6}$$

In the first case, there is no intensity of development $n(x)$ that abides by the zoning constraint and that delivers utility of at least $\bar{u}$. This could happen because $\bar{n}(x) < n_L(x)$ or because $OL\,(n, x) < MC\,(n)$ for every $n(x) > 0$. In the second case, the zoning constraint binds and the equilibrium price is consistent with household spatial equilibrium. In the third case, the zoning constraint does not bind and the equilibrium price is consistent with both household spatial equilibrium and (unconstrained) profit maximization.

1.2.1.3 Locally Optimal Zoning Choice

This section describes the determination of zoning laws within a location.

I assume that some location have a measure 0 of *initial residents* who choose the zoning regulation, while some locations have no initial resident and are thus unregulated. The set inhabited locations will be endogenized later. Formally, they can choose the zoning constraint $\bar{n}(x)$; as Equation (1.6) makes clear, the choice of zoning law affects the local equilibrium.

[19] See appendix for details on equilibrium stability.

For the subset of locations with no initial residents, the zoning law is effectively infinite: $\bar{n}(x) = \infty$. The equilibrium population in such locations is given by Equation (1.6): it will be $\{n_H(x), MC(n)\}$ if the $OL(n,x)$ and $MC(n)$ curves intersect and $\{0,0\}$ if the $OL(n,x)$ curve is strictly lower than the $MC(n)$ curve for all positive populations.

Locations with initial residents The measure 0 of initial residents have identical preferences and productivity as the households described above, but they differ in two respects. First, the initial residents of location x face a fixed rent $p_0(x)$. I take this rent as given for now, but will endogenize it later. Second, they choose the zoning constraint $\bar{n}(x)$ that limits the maximum level of development within their location. The initial residents have rational expectations about the consequences of their zoning choice. Namely, they anticipate that a zoning constraint $\bar{n}(x)$ will induce the local spatial equilibrium of Equation (1.6). Like the households described above, the initial residents take as given the the endogenous outside option $\bar{u}$ and the level aggregate profits Π.

Let $\theta(\bar{n}(x); \bar{u}, \Pi) = n^*(x)$ denote the population $n^*(x)$ from Equation (1.6) given a zoning constraint $\bar{n}(x)$ and aggregate variables $\bar{u}$ and Π. The initial residents solve the following maximization problem:

$$\max_{\{c_0(x), \bar{n}(x)\}} u(c_0(x)) - v(\theta(\bar{n}(x); \bar{u}, \Pi))$$

subject to

$$c_0(x) = y(x) + \Pi - \frac{F}{\theta(\bar{n}(x); \bar{u}, \Pi)} - p_0(x)$$

We can rewrite the maximization problem in a more convenient form. Initial residents in location x act as if they were directly choosing the local population given $\bar{u}$ and Π, subject to a household participation constraint.

Proposition 2. *Consider an arbitrary x for which there exists n with $OL(n,x) \geq$*

$MC(n)$. For this x, the locally optimal zoning constraint solves:

$$\max_{\{c_0(x),\bar{n}(x)\}} u(c_0(x)) - v(\bar{n}(x)) \tag{1.7}$$

subject to the budget constraint

$$p_0(x) + c_0(x) + \frac{F}{\bar{n}(x)} = y(x) + \Pi \tag{1.8}$$

and the participation constraint

$$u\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - 2\bar{n}(x)/Z^2\right) - v\left(\bar{n}(x)\right) \geq \bar{u}. \tag{1.9}$$

Proof. See appendix. □

Note that the participation constraint is a function of the marginal cost of housing, rather than the equilibrium price. The participation constraint is satisfied if and only if $\bar{n}(x) \in [n_L(x), n_H(x)]$ as defined above. If the participation constraint is non-binding, Equation (1.6) states that equilibrium population $n^*(x)$ will equal the initial resident choice $\bar{n}(x)$. The logic of the constraint is as follows: for $\bar{n}(x) < n_L(x)$, the equilibrium population will be 0. However, the initial residents prefer to share the fixed cost F with a positive population, and so they will not choose $\bar{n}(x) < n_L(x)$. If they choose $\bar{n}(x) > n_H(x)$, the equilibrium population will be $n_H(x)$, and so limiting the initial resident choice to being below $n_H(x)$ does not restrict their potential payoffs.

In general, the participation constraint may or may not bind. If the participation constraint does not bind, then the zoning constraint will bind in the local equilibrium: the equilibrium house price is above the firm's marginal cost. If the participation constraint does bind, then the zoning constraint does not bind: the house price is equal to marginal cost.

When the participation constraint does not bind,[20] the initial resident choice of zoning choice is the $\bar{n}(x)$ that solves

$$u'\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - p_0(x)\right)\frac{F}{\bar{n}(x)^2} - v'\left(\bar{n}(x)\right) = 0. \tag{1.10}$$

The first term is the marginal benefit of sharing the fixed cost F more broadly. The second term is the marginal cost of congestion. As u is concave and v convex, the expression is strictly decreasing and the $\bar{n}(x)$ that solves the equation is unique.[21]

When the participation constraint binds, the initial resident's optimal zoning choice is by definition not in the interior of the set of points consistent with the participation constraint. By concavity of the initial resident problem, the zoning choice in these cases is given by the endpoint of the set that offers greater utility to the initial residents:

$$\underset{\bar{n}(x)\in\{n_L(x),n_H(x)\}}{\arg\max}\; u\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - p_0(x)\right) - v\left(\bar{n}(x)\right). \tag{1.11}$$

1.2.2 General Spatial Equilibrium

Define a *stable general spatial equilibrium* in this economy to be an endogenous outside option $\bar{u}$, a level of profit Π, a set of open locations $\mathcal{X}$, local populations $n^*(x)$, local prices $p^*(x)$, and zoning laws $\bar{n}(x)$ such that the zoning choices $\bar{n}(x)$ and local equilibrium outcomes $\{p^*(x), n^*(x)\}$ solve household, firm, and initial resident problems and are

[20] The constraint states that the utility delivered by the location *with house price equal to marginal cost* is greater or equal to $\bar{u}$. If the constraint is non-binding then Equation (1.6) states that the equilibrium house price will be greater than marginal cost.

[21] As there is a measure 0 of initial residents, an inflow of $\bar{n}(x)$ households ensures that the new households will be the majority of the community. This raises the question of whether they would seek *ex post* to hold a new vote and modify the zoning law. However, the new residents would not have a strong preference to modify the law. Given an outside option $\bar{u}$, households expect that rents will adjust to make them indifferent across locations regardless of their vote. As such, households do not strictly prefer any alternative zoning law for their location. The pre-fixed rent of the initial residents eliminates this feedback mechanism.

consistent with the population constraint

$$\int_0^1 n^*(x)dx = 1, \tag{1.12}$$

with the definition of profits

$$\Pi = \int_0^1 \left[p^*(x)n^*(x) - \left(\frac{n^*(x)}{Z} \right)^2 \right] dx, \tag{1.13}$$

with the definition of $\mathcal{X}$

$$x \in \mathcal{X} \iff n^*(x) > 0, \tag{1.14}$$

and with the stability condition

$$p^*(x) = p_0(x) \ \forall \ x \in [0, 1] \, . \tag{1.15}$$

The stability condition imposes restrictions on the prices faced by initial residents. This natural restriction is informed by a notion of dynamic stability and enables the static model to approximate the steady-state of a corresponding dynamic overlapping generations model, wherein a subset of agents inherit a fixed rent from the previous period's equilibrium.

1.2.2.1 Characterization of General Equilibrium

I proceed by characterizing the triplet $\{\bar{u}, \Pi, \mathcal{X}\}$ as well as optimal zoning $\bar{n}(x)$, local population $n^*(x)$, and price gradient $p^*(x)$ consistent with general equilibrium. Focusing on the *stable equilibrium* described above, I consider the natural restriction that $p_0(x) = p(x)$. Two preliminary results will be useful in characterizing the general equilibrium.

Proposition 3. *The set of occupied locations $\mathcal{X}$ is $[\underline{x}, 1]$ for some threshold location $\underline{x}$. For the threshold location $\underline{x}$, the $OL(n, \underline{x})$ and $MC(n)$ curves are tangent at a unique* *level of population that depends on $\bar{u}$.*

Proof. See appendix. □

This tangency condition proves useful for pinning down the price and population gradients. The intuition for tangency is as follows: from Proposition 1, the two curves cross twice, once with tangency, or not at all. For a given location x, if they do not cross the location must be unoccupied: no level of population can deliver utility $\bar{u}$. If they do cross or meet with tangency, then $n^*(x)$ must be strictly positive. Otherwise, there would be no initial residents, no zoning constraint, and developers would find it optimal to build (and households to settle) to the local equilibrium population of $n_H(x)$. Thus x must be in $\mathcal{X}$ for all locations where the $OL(n, x)$ curve meets the $MC(n)$.

The tangency of $OL(n, \underline{x})$ and $MC(n)$ at the threshold location $\underline{x}$ also implies equality between $OL(n, \underline{x})$ and $MC(n)$. Thus Proposition 1 offers two additional conditions that will be used to pin down the price and population gradients. First, tangency implies

$$\frac{F}{n^*(\underline{x})^2} - \left[u^{-1}\right]' \left(v\left(n^*(\underline{x})\right) + \bar{u}\right) \times v'\left(n^*(\underline{x})\right) = 2/Z^2. \tag{1.16}$$

Second, equality implies

$$p(\underline{x}) = 2n(\underline{x})/Z^2. \tag{1.17}$$

The first condition pins down the threshold population consistent as a function of $\bar{u}$. The second condition pins down the price at the threshold as a function of this population.

Lemma 1. *Define $\bar{n}$ to be the population that maximizes the $OL\,(n, x)$ curve. Given outside option $\bar{u}$, the level of population $\bar{n}$ is unique and independent of location.*

Proof. See Appendix. □

Lemma 2. *Define location* $\hat{x}(\bar{u}, \Pi)$ *as follows:*

$$\hat{x}(\bar{u}, \Pi) = \begin{cases} x \text{ s.t. } OL(\mathcal{N}(\bar{u}), \hat{x}) = MC(\mathcal{N}(\bar{u})) & \text{if } OL(\mathcal{N}(\bar{u}), 1) \geq MC(\mathcal{N}(\bar{u})) \\ 1 & \text{otherwise.} \end{cases}$$

Then $\hat{x}(\bar{u}, \Pi)$ *is unique and in the range* $[\underline{x}, 1]$.

Proof. See appendix. □

Proposition 4. *In a stable general equilibrium, the following characterizes the local equilibrium populations and optimal zoning. For all locations* $x > \hat{x}$, $n^*(x) = \bar{n}(x) = \mathcal{N}(\bar{u})$. *For all locations* $x \in [\underline{x}, \hat{x}]$, *the population* $n^*(x)$ *and optimal zoning* $\bar{n}(x)$ *are given by the upper intersection of the* $OL(n, x)$ *and* $MC(n)$ *curves. For all* $x \geq \underline{x}$, $p^*(x)$ *is given by the value of the* $OL(n, x)$ *curve evaluated at* $n^*(x)$.

Proof. See appendix. □

The $OL(n, x)$ curve reflects the willingness to pay for a location as a function of its population: if a mix of congestion and sharing offers a higher level of utility, households are willing to pay more. As $p_0(x) = p(x)$ in the stable equilibrium, initial resident utility is maximized where the willingness to pay is maximized: at the peak of the $OL(n, x)$ curve. As the productivity $y(x)$ only shifts the $OL(n, x)$ curve vertically, the preferred zoning choice is independent of location. Hence for locations with $x > \hat{x}$, zoning laws bind in equilibrium at the population $\mathcal{N}(\bar{u})$.

For locations $x \leq \hat{x}$, the level of population $\mathcal{N}(\bar{u})$ preferred by the initial residents is inconsistent with local equilibrium: the price that would induce households to settle at population $\mathcal{N}(\bar{u})$ is too low to induce construction firms to build $\mathcal{N}(\bar{u})$ units of housing. Vice versa, any price high enough to induce firms to build $\mathcal{N}(\bar{u})$ will be higher than households are willing to pay, given local wages $y(x)$. In this case, the equilibrium population adjusts downwards: the house price falls, marginal costs fall, and so long as $x \geq \underline{x}$, the

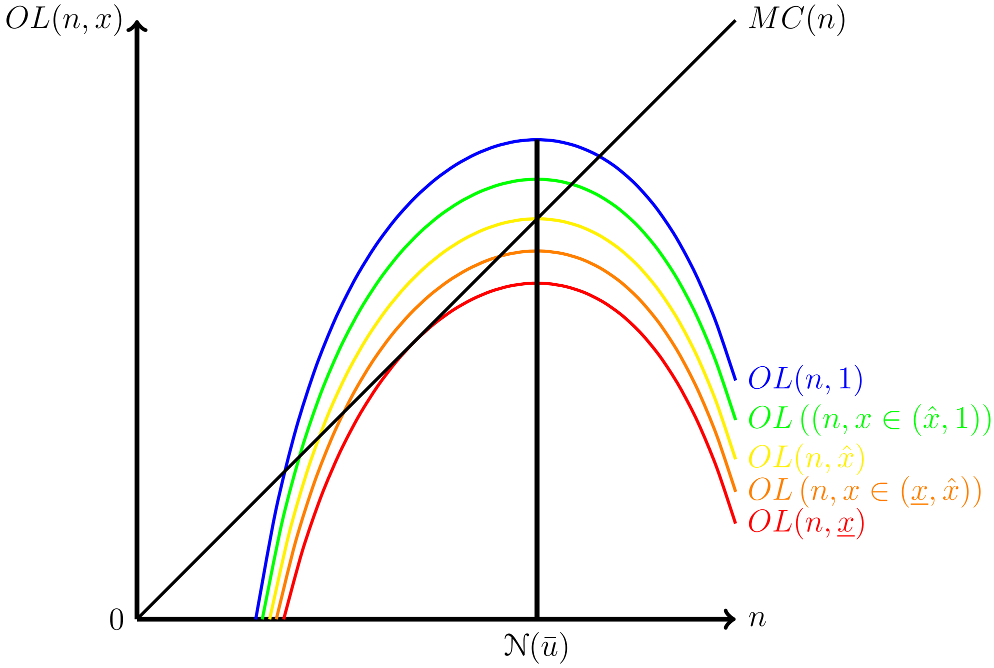

Figure 1.2: Optimal location $OL(n,x)$ curves for threshold occupied location $\underline{x}$, $\hat{x}$, and other locations with the marginal cost $MC(n)$ curve. The lowest curve, $OL(n,\underline{x})$, is the threshold location and its equilibrium population will be given by the point of tangency with the $MC(n)$ curve. For the next curve, the equilibrium population is given by the intersection on the right. For the top three curves, the initial resident zoning choice and equilibrium population is given by $\mathcal{N}(\bar{u})$. Note that all of the $OL(n,x)$ curves reach their peak at $\mathcal{N}(\bar{u})$; the equilibrium population at each location is as close to this as possible subject to ensuring that $OL(n,x) \geq MC(n)$.

location settles at a smaller equilibrium population. Because the initial resident problem is concave, the participation constraint binds and the local equilibrium population (and optimal zoning law) is the upper intersection of the $OL\,(n,x)$ and $MC\,(n)$ curves. This case is shown in Figure 1.2.

Characterization of aggregate variables: $\bar{u}$ and Π. As noted, the equilibrium levels of population $n^*(x)$ and prices $p^*(x)$ described in Proposition 4 are consistent with the local equilibrium, optimal zoning, and stable equilibrium conditions for a given pair $\{\bar{u},\Pi\}$. The general equilibrium $\bar{u}$ and Π are the pair for which the local equilibria $n^*(x)$ and $p^*(x)$ and locations $\underline{x}$ and $\hat{x}$ described above are consistent with the population constraint and profit definition.

The threshold conditions from Proposition 3 implicitly define the threshold location $\underline{x}$ as a function of $\bar{u}$ and Π: $\underline{x}(\bar{u},\Pi)$. The profit definitions states

$$\Pi = \int_{\mathcal{X}} \left(p^*(x)n^*(x) - \left(\frac{n^*(x)}{Z}\right)^2 \right) dx.$$

Substituting the equilibrium conditions, this becomes

$$\begin{aligned}\Pi = &\left(\int_{\hat{x}(\bar{u},\Pi)}^{1} y(x)dx + \Pi - \frac{F}{\mathcal{N}\,(\bar{u})} - u^{-1}\left(v(\mathcal{N}\,(\bar{u})) + \bar{u}\right) \right) \mathcal{N}\,(\bar{u}) \\ &+ \int_{\underline{x}(\bar{u},\Pi)}^{\hat{x}(\bar{u},\Pi)} \left(\frac{n_H(x)}{Z}\right)^2 dx - \left(\frac{\mathcal{N}\,(\bar{u})}{Z}\right)^2.\end{aligned} \tag{1.18}$$

From above, $n_H(x)$ is the upper intersection of the $OL(n,x)$ and $MC(n)$ curves and is pinned down as a function of Π and $\bar{u}$. From the population constraint,

$$1 = \int_{\underline{x}(\bar{u},\Pi)}^{\hat{x}(\bar{u},\Pi)} n_H(x)dx + \int_{\hat{x}(\bar{u},\Pi)}^{1} \mathcal{N}\,(\bar{u})\,dx. \tag{1.19}$$

The characterization of a stable general equilibrium is completed by an outside option $\bar{u}$ and a profit Π that solve Equations (**??**) and (1.19). I do not yet have a formal existence proof. However, given functional forms for $u(c)$ and $v(n)$, it is straightforward to solve the equations and characterize the equilibrium numerically.

1.2.2.2 Qualitative Predictions

The model can be used to generate a set of empirical predictions regarding density, house prices, and marginal costs. Regarding density, the model predicts that the equilibrium density will be uncorrelated with local productivity for locations where zoning is binding. For all tracts, it should be positive. For the set of urban census tracts, the correlation between income and density is -0.02. Adjusting for skill differences, in a process to be detailed below, moves the correlation to 0.02. After adjusting for county-level averages—which may result from common investments in infrastructure, or exogenous amenities—the correlation moves to 0.07. These findings are consistent with the model.

The model predicts that house prices will perfectly offset productivity differences in locations with binding zoning. Of course, housing prices respond to many factors not included in the model, so the data is unlikely to show a perfect fit. To test this prediction, I compute a residual house price by regressing tract-level housing costs on the same observable characteristics that I use to adjust income. Then, I correlate this measure with tract-level income, adjusted for observable characteristics. The correlation is 0.42 for owner-occupied housing costs, and 0.38 for rental costs. These correlations are not inconsistent with the model. In addition to the aforementioned exogenous amenities, unmodeled forces that may affect this relationship include the quality of the housing stock, length of ownership for homeowners, and housing subsidies or rent-control laws.

Finally, the model predicts that marginal costs will be constant across locations with binding zoning. Together, these last two predictions imply that house prices in the most expensive locations will be well above marginal costs of construction. This prediction is well-supported by the (cf. Glaeser et al., 2005a).

1.2.3 Optimality: The Constrained Planner's Problem

To provide a welfare benchmark with which to contrast the zoning equilibrium this section describes the problem of a constrained planner. The planner chooses the set $\mathcal{X}$ of locations to open, allocates population $n(x)$ among these locations, and allocates consumption $c(x)$ to households in order to maximize welfare. The planner is free to transfer output across locations. The planner faces standard population and aggregate resource constraints as well as a spatial equilibrium constraint. This final constraint restricts the planner to choosing allocations that deliver identical utility to each household regardless of location; in this sense, the planner is *constrained*. The spatial equilibrium condition implies that the welfare criterion to be maximized is simply $\bar{u}$. The set of locations to open can be simplified to the choice of a threshold location $\underline{x}$.[22]

The planner's problem is thus given by the following:

$$\max_{\{\bar{u},c(x),n(x),\underline{x}\}} \bar{u}$$

subject to the spatial equilibrium constraint

$$u\left(c(x)\right) - v\left(n(x)\right) = \bar{u} \ \forall \ x \in \left[\underline{x}, 1\right],$$

the population constraint

$$\int_{x=\underline{x}}^{1} n(x)dx = 1,$$

and the aggregate resource constraint

$$\int_{x=\underline{x}}^{1} \left[\left(y(x) - c(x)\right) n(x) - F - \left(\frac{n(x)}{Z}\right)^2\right] dx = 0.$$

[22] The planner will never choose to open location x_1 if location $x_2 > x_1$ has not already been opened. Hence choosing the set $\mathcal{X}$ from within the set of locations $[0, 1]$ amounts choosing the lowest-productivity location to open: $\underline{x}$.

The population $n(x)$ can be thought of as the *intensive* margin of development while $\underline{x}$ is the *extensive* margin of development. Note that each opened location is subject to a spatial equilibrium condition. Given the choice of intensive margin $n(x)$, the location-specific spatial equilibrium constraints ensure the planner engages in transfers of output such that household consumption offsets the level of congestion and each household receives utility $\bar{u}$.

The Lagrangian associated with the problem is:

$$\begin{aligned}\mathcal{L} = \bar{u} + \int_{x=\underline{x}}^{1} \lambda(x)\left[u\left(c(x)\right) - v\left(n(x)\right) - \bar{u}\right] dx + \mu \left[1 - \int_{x=\underline{x}}^{1} n(x) dx\right] \\ + \Lambda \int_{x=\underline{x}}^{1} \left[\left(y(x) - c(x)\right) n(x) - F - \left(\frac{n(x)}{Z}\right)^2\right] dx.\end{aligned} \tag{1.20}$$

Here, $\lambda(x)$ is the Lagrange multiplier on the spatial equilibrium constraint at location x, μ is the multiplier on the population constraint and Λ is the multiplier on the resource constraint. Taking first order conditions for $n(x)$ and $c(x)$ and rearranging, the planner weighs the following objects against one another when choosing the intensive margin of development $n(x)$:

$$\Lambda\left[y(x) - c(x) - 2n(x)/Z^2\right] - \Lambda n(x)\frac{v'(n(x))}{u'(c(x))} - \mu = 0 \tag{1.21}$$

The first term is the shadow resource value of the marginal household, the second term is the shadow cost of congestion, in terms of resources, and the third term is the shadow cost of having one fewer household to allocate elsewhere. The marginal household in a location adds output—net of consumption and the construction cost—but also increases congestion for the $n(x)$ households already there.

To recall, the first order condition for the choice of local zoning was

$$\frac{F}{n(x)} - n(x)\frac{v'(n(x))}{u'(c(x))} = 0.$$

Both the planner and the initial residents consider the role of the congestion externality. While the initial residents weigh this externality against the value of sharing the local fixed cost more broadly, the planner weighs it against the resource value of allocating a marginal household to location x and the shadow value of the binding population constraint. The first term in Equation (1.21) represents the net output of allocating a marginal household to location x. Under the flat gradient that arises from the local zoning equilibrium, the marginal value would be higher in productive locations. The planner will thus allocate more households to such locations. This is they key intensive-margin wedge between the optimal solution and the allocation with local zoning.

While initial residents ignore aggregate effects, they do consider the marginal value of sharing the fixed cost more broadly. The planner ignores this margin: the fixed cost is paid when the location was opened and should not affect the intensive margin. This is the second wedge between the optimal solution and the allocation with local zoning. In short, the initial residents ignore the aggregate effects of their choice to restrict housing supply.

The planner instead considers the fixed cost F at the extensive margin. As shown in Equation (1.22), the planner weighs the net output of opening the marginal location $\underline{x}$ against the shadow value of assigning $n(\underline{x})$ households to this location.

$$\Lambda \left[n(\underline{x}) \left(y(\underline{x}) - c(\underline{x}) \right) - \left(\frac{n(\underline{x})}{Z} \right)^2 - F \right] - n(\underline{x})\mu = 0. \tag{1.22}$$

Recall that in the local zoning allocation outlined previously, the location in which the $OL(n, x)$ and $MC(n)$ are tangent becomes the threshold. Restating, this condition was met where

$$\frac{F}{n^*(\underline{x})^2} - \left[u^{-1} \right]' \left(v \left(n^*(\underline{x}) \right) + \bar{u} \right) \times v' \left(n^*(\underline{x}) \right) = 2/Z^2.$$

This location arises through the general spatial equilibrium, rather than through being chosen directly by any agent. In choosing the optimal threshold, the planner weighs the value and costs of opening a marginal location. The differential outcomes for the threshold $\underline{x}$ highlight an additional externality generated by the initial resident choice of local zoning: restrictive zoning ensures that too many locations are opened in equilibrium, necessitating the payment of fixed costs to open locations that would not be paid under the optimal allocation.

1.2.3.1 Decentralization: Socially Optimal Zoning

The planner's allocation can be decentralized through the adoption of an alternative zoning regime. In this decentralization, the full measure 1 of households votes on the set $\mathcal{X}$ of locations to open and the gradient of zoning laws $\bar{n}(x)$ for each location x in $\mathcal{X}$. In so doing, households have rational expectations over the spatial equilibrium induced by the choices they make. Households therefore choose zoning laws $\bar{n}(x)$ and the set of opened locations $\mathcal{X}$ to maximize the endogenous outside option $\bar{u}$. Households choose these laws subject to the population constraint and to the local and general spatial equilibria conditions.

Proposition 5. *The utility-maximizing choices for the set of opened locations $\mathcal{X}$ and the set of zoning laws $\bar{n}(x)$ are identical to those chosen by the planner. The equilibrium price gradient and aggregate profits implement the planner's choice of consumption. This set of instruments allows households to fully implement the planner's allocation.*

Proof. See appendix. □

Intuitively, households have the option of choosing the same zoning as the planner and will do so. Prices will adjust such that households are indifferent across all locations. With identical population across locations as the planner's allocation, total output and aggregate construction and fixed costs are also identical. By resource balance, aggregate consumption must also be identical. By spatial equilibrium, consumption must make households indifferent between locations, and therefore equilibrium consumption is identical to the planner's allocation. The zoning laws and set of opened locations chosen by households in this problem can therefore be described as *socially optimal zoning*, as opposed to the *locally optimal zoning* described previously.

The profits paid to households are a crucial mechanism for ensuring that this spatial equilibrium will also be consistent with construction firm behavior. Note that household spatial equilibrium pins down relative prices, but not levels. Because profits are distributed to households regardless of location, an increase of ε in the absolute level of prices at each location leads to an increase in household profits by ε. This increase leaves household consumption identical. This mechanism ensures that there exists a price gradient consistent with local equilibrium from the firm perspective.

1.3 Calibration

1.3.1 Data

Four key model components must be calibrated to data: the magnitude of the local fixed cost, the shape of utility from consumption, the disutility of congestion, and the distribution of location-specific productivity. This section describes the moments in the data that the calibration seeks to match. In matching the moments, I take the set of urban U.S. Census tracts to be the empirical counterpart of model locations.

For the utility of consumption, I consider $u(c) = c$. Linear utility follows related papers (Van Nieuwerburgh and Weill, 2010; Davis and Dingel, 2014).[23]

For the disutility of congestion, I consider $v(n) = \gamma n^{\eta}$. To calibrate γ, I match the fraction of income spent on consumption goods as measured in the Consumer Expenditure Survey of the Bureau of Labor Statistics. As γ controls the relative utility weights of consumption and congestion, a higher γ will induce households to open more locations and spend more on housing and fixed costs rather than consumption. To ensure the

[23]See appendix for robustness with log utility.

data is well-matched to the model concept, calibrate the consumption share of income to spending on tradeable goods. These spending shares range from 52–61% for different income deciles, and average 58% overall.

As noted previously, an alternate interpretation of the local fixed cost has residents gaining utility from a greater diversity of monopolistically competitive local services, each of which is subject to a fixed cost.[24] Using this interpretation, I calibrate the fixed cost F to match the share of consumer spending on local services. In particular, I use the Consumer Expenditure Survey categories for food away from home, personal services, medical services, and fees and admissions. These spending shares range from 7–10% for different income deciles, and average 8.5% overall.

I calibrate the congestion shape parameter η to match the average population density of urban census tracts, as identified by the Census Bureau for 2010. Given the functional form assumptions, the local equilibrium population for locations with binding zoning is given by

$$n = \left(\frac{F}{\gamma\ \eta}\right)^{1/(1+\eta)}.$$

As the zoning law will bind for a substantial fraction of locations, the shape parameter η has a first-order effect on average population density.[25] Intuitively, too low a congestion cost would cause households to live at too-high densities in productive locations, compared to the data.[26]

I calibrate construction firm productivity to match the ratio of price per square foot to cost per square foot in the most productive places: Manhattan, San Francisco, or

[24] In a local services model, expenditure will be precisely $F/n(x)$ if the elasticity of substitution between local service firms is 1 or if aggregated local services enter the utility function linearly, such that $u(c, S) = c + S$, where S is a CES aggregator of local services. More generally, local expenditure will be $F/n(x)^\sigma$, where σ is the elasticity of substitution between local services.

[25] The intuition is similar in spirit to using wage shares to calibrate the labor and capital exponents in a production function: in equilibrium, the parameters can have first-order effects on levels.

[26] Note that it is not necessary to take a stand on the precise productivity of *non-urban* locations, whose measured incomes may not be a good guide to those earned by new households. It is sufficient to assume that non-urban households have lower productivity than those occupied at urban population levels.

Param	Value	Description	Model Target	Data	Model
γ	1/215	Congestion weight	Consumption share of expenditure	52%–61%	55%
F	0.1265	Fixed cost	Local services share of expenditure	7%-10%	7.8%
η	7	Congestion curvature	Average urban density	5,100 per mi sq mi	5,900 per sq mi
Z	2.7	Construction productivity	p/MC in best location	5.2–5.8	5.6

Table 1.1: Calibration. The classification of tracts as *urban* follows the Census Bureau's 2010 classification. The consumption share of expenditure is taken from the Consumer Expenditure Survey for 2012, and is equal to consumption net of spending on housing and local services. The price-to-marginal cost ratio follows the approach of Glaeser et al. (2005a).

Silicon Valley. For this exercise, average price per square foot data is taken from Zillow, and cost data is taken from RS Means. Both report cost data at the level of counties and, in some instances, for sub-county regions.

The distribution of income is taken from census tract median incomes, adjusted for observables in a process that follows Hsieh and Moretti (2015). I take individual-level data from IPUMS on income, education, race, and gender. I then run a regression using the following equation:

$$y_i = X_i\beta + \varepsilon_i,$$

where y_i is the income of individual i and X_i is the vector of observable characteristics. I then take the estimate of β and calculate the residual income $\tilde{y}$ of each tract ℓ:

$$\tilde{y}_\ell = y_\ell - X_\ell\beta,$$

where y_ℓ is the measured average income per worker and X_ℓ is the fraction of the tract with the given observable characteristics.

Table 1.3.1 summarizes the calibration targets and the model fit. For three of the four parameters, the calibrated parameter is within the range from the data. The search of the parameter space failed to find a set of values for which the density was closer to the data while maintaining consistency with the other parameters. At the same time, it is plausible that households have reasons beyond productivity for choosing locations.

1.4 Quantitative Results

The key quantitative question this paper addresses is: what are the welfare costs of locally-determined zoning laws? In the baseline calibration, implementing the planner's allocation would increase welfare by 1.4%. Consumption would increase by 2.4% and GDP by 2.1%, but increased congestion would mitigate these gains.[27] The planner would close the 3% lowest-productivity locations, and population density at the 95th percentile location would increase by 18%. The density at the median location would increase by 2%. These results are summarized in Table 1.2, along with results from two additional counterfactuals.

The first counterfactual, labeled *Market* in the table, corresponds to an allocation with no zoning, and construction firms free to build without constraint. In the language of the model, the equilibrium for each location is the upper intersection of the $OL(n, x)$ and $MC(x)$ curves. This counterfactual corresponds to that studied by Hsieh and Moretti (2015), and I find an increase in GDP of 6%, half of the 13.5% that they report. Part of the different may be due to the presence in this model of congestion externalities, which limit the willingness of households to crowd into productive locations. Moreover, the estimates are of the same order of magnitude despite using different methodologies.

While the GDP increase from zoning abolition is substantial, the increased in congestion is as well. The median location sees an increase in population density of 15%, and the increase is almost 50% at the 95th percentile. Accordingly, welfare declines by almost 6%, despite the large increase in consumption. This result suggests that the productivity

[27]Recall that utility is quasilinear, so the welfare and consumption figures can be compared meaningfully.

Relative Outcome	Planner	Market	Upzoning
Welfare	1.4%	-5.9%	-2.0%
GDP	2.1%	6.0%	2.0%
Consumption	2.4%	5.6%	2.1%
Median Rent	2.9%	-16%	-1.8%

Table 1.2: Main Results. All results are relative to the main zoning results. The planner results give the gains from moving to the optimal allocation. The market results give the counterfactual with no zoning laws. The upzoning results give the counterfactual with no zoning laws in the 5% highest-productivity locations.

gains to zoning abolition put forth by Hsieh and Moretti (2015) will not, in fact, increase welfare.

The second counterfactual, labeld *Upzoning* in the table, corresponds to an allocation with no zoning in locations with the productivity greater than the 95th percentile. As census tracts average about four thousand inhabitants, these locations are home to approximately ten million residents. For scale, the San Francisco Bay Area is home to five million urban residents. Under this allocation, GDP increases by 2% while welfare declines by 2%. This contrasts with Hsieh and Moretti (2015), who find large productivity gains of 9.5% from increasing the population of the best cities.

The median rent, also in Table 1.2, provides insight into the changing tradeoffs with each reform. Three key forces determine the median rent: the productivity of the median resident, the congestion faced by the median resident, and the outside option $\bar{u}$. Under the decentralized version of the planner's allocation, house prices *increase* by 3% for the median household. Each location is a little more crowded—and so rents for each location fall, between 1–5%—but the median household is now in a more productive location and thus their rent increases. The market equilibrium also sees the median household in a more productive location, but this location is now much more crowded. Correspondingly, rents fall by 16%. Finally, in the upzoning case, the median household is in a more productive location with the same level of congestion. However, they are slightly worse off, and so they are unwilling to bid as much for housing, and rents fall.

Given the calibrated parameters, the main policy implication of the model is to allow zoning laws to be chosen at a higher geographic level—preferably, the national level. This would ensure that the productivity effects of zoning laws are internalized and would enable the implementation of the planner's optimal allocation.

The fundamental role of congestion in shaping preferences and outcomes points towards a second policy implication beyond implementing the planner's regime: introducing reforms that could lower the cost of congestion to the neighborhood. Understanding the specific component of congestion that drives externalities will identify the appropriate form of mitigation. If neighborhood congestion is driven by vehicle traffic, then perhaps rapid transit would enable more-dense development by reducing the costs of congestion imposed by new development. Congestion-mitigation efforts like transit can be quite costly, and within a context of spatial equilibrium, the benefits would be felt widely. As with local zoning, the current transit-planning regime may not take into consideration the external effects of transit.

Distributional Outcomes Abolishing zoning—whether nationally or just in the highest-productivity locations—cannot improve welfare under this calibration. This fact is driven in part by the curvature to the cost of congestion: increasing the intensity of development in productive places is simply too costly in terms of welfare. A second factor that plays an important role is the assumption that profits are shared equally among all households. When restrictive zoning drives up house prices in productive locations, the profits earned are redistributed across all locations. To understand the role played by this assumption, I now introduce a second calibration wherein construction-sector profits are not returned to households. This calibration mirrors the traditional urban economics assumption of *absentee landlords*, who collect rent but do not otherwise interact with households.[28]

[28]Eeckhout and Guner (2015) perform a similar exercise and identifies absentee landlords with the concentrated landholdings of the 1% wealthiest households. They suppose that the planner values a fraction of housing sector profits that is distributed to households, but not the fraction distributed to absentee landlords.

Relative Outcome	Planner	Market	Upzoning
Welfare	1.5%	-10%	-3.2%
Welfare ex profits	7.3%	23%	3.7%
GDP	2.6%	6.3%	1.9%
Consumption	9.3%	50%	12%
Median Rent	-5.3%	-42%	-2.1%

Table 1.3: Distributional Outcomes. All results are relative to the main zoning results, and all results ignore the welfare derived from profits—this allows the welfare gains *discounting profits* from the market allocation to exceed those from the planner's. The planner results give the gains from moving to the optimal allocation. The market results give the counterfactual with no zoning laws. The upzoning results give the counterfactual with no zoning laws in the 5% highest-productivity locations.

Table 1.3 shows the analogous results under this model.[29] Here, the gains to GDP are similar to those before, and the losses to welfare of zoning abolition are even greater. However, the change to welfare discounting profits—that is, the welfare of the typical households—are positive. This exercise highlights the key role played by construction profits. Residents of productive places don't actually enjoy the high productivity, they simply pay higher rents. When these profits are shared broadly, these rents are redistributed so that all households, regardless of location, share the output of high-productivity locations. When these profits are not shared broadly, households in less productive locations have more to gain from zoning abolition because moving into a productive location now earns them significantly more consumption.

1.4.1 Empirical Implications

The model focuses on just two of potentially many local externalities. Numerous facets of the model lend themselves to further empirical validation. Similarly, the simple model can be extended with greater heterogeneity, including of household productivity, of locational or household preferences over congestion costs and fixed costs, and of locational amenities.

[29]For these results, I have recalibrated the model so that the new model-defined moments are consistent with the data. The consumption share is 54%, the local service share 7.7%, the construction sector markup is 7.4%, and 88% of locations in the data are opened.

Within the model, existing residents choose zoning regulations to maximize their amenity mix, trading sharing externalities against congestion. In doing so, they also maximize the rents that new households are willing to pay: maximizing amenities and maximizing land values are identical. However, local ownership of land (e.g., in the form of homeownership) may break this felicitous duality as landowners may seek to restrict the supply of new housing, a close substitute to their own asset. Empirically, it may be possible to distinguish between these intents using evidence from surveys or by identifying preferences over regulations that would increase supply but have a positive effect on local amenities (or vice versa).

Local variation in model parameters could be tested to examine the strength of their relationship to equilibrium outcomes. For instance, the fixed costs of development may be higher in regions that are more arid, or prone to extreme weather, due to the costs of necessary infrastructure. New development in the arid fringe of Los Angeles is quite dense by the standards of new development in rainier eastern cities like Atlanta.[30] All else equal, do locations with higher fixed costs see higher-density development, as the model would predict?

Similarly, some locations have seen dramatic changes to their productivity. Homeowners in these locations—like the initial residents of the model—don't see any immediate change in their costs. The model predicts that locations with large positive shocks should see little new development but great increases in house price. Instead, the model predicts that new development would be concentrated in previously-rural locations at the fringes of newly productive regions. Anecdotally, Palo Alto—a key center of innovation in the San Francisco area—has grown from a population of 55,225 in 1980 to 64,403 in 2010. The city of Antioch, of similar geographic extent at the fringe of the region, grew from 43,559 residents to 102,372 over the same period. Part of this differential can be explained

[30]For example, a recently-developed subdivision in Lancaster, CA was developed to approximately 5,000 people per square mile, while a similarly-new development in Powder Springs, GA is approximately half the density.

by the effects of durable capital (cf. Siodla, 2015); a careful test of this prediction could be challenging but insightful.

In Boulder, CO, a measure on the ballot in the fall of 2015 would have empowered dozens of local neighborhood groups with an easy path to halting unwanted development, within a city of just 100,000 residents. The outcome and voting patterns in the election could help to test model predictions. In particular, the model assumes locations are sufficiently small that residents can ignore the aggregate effects of their zoning decisions. Regulatory changes at higher levels of government are more likely to affect aggregate variables, such as the threshold occupied location. Residents who stand to benefits more from changes to aggregate variables would be more likely to oppose the measure. Renters, the young, and new arrivals may be more likely to benefit from aggregate changes within Boulder, relative to long-time homeowners with substantial equity dependent on the status quo.

Similarly, the model could be extended to include differential preferences over congestion: perhaps some households are more or less bothered by adding more neighbors. Empirically, households with families might be more sensitive to congestion, even conditioning on dwelling size. This extension could refine the estimates of welfare gains from adding density to productive locations.

1.5 Conclusion

This paper provides a unified framework to address the local and aggregate welfare effects of local land-use regulation. It provides empirically-grounded externalities that incent households to pass restrictive zoning laws that prevent new housing development at the neighborhood level. In their endogenous choice of zoning, households rationally ignore the aggregate implications that arise from location heterogeneity. Households thus bid up the price of artificially scarce housing in productive locations so that prices exceed the marginal costs of construction.

The endogeneity of zoning plays a critical role in overcoming the lack of sufficient data to identify the wedge between house price and marginal cost at the neighborhood level. A calibrated version of the model is used to perform a welfare calculation: how large are the aggregate losses to welfare from local zoning relative to the planner's optimum? The model provides an interpretation of these welfare costs without measuring these wedges directly. In the preferred calibration, welfare could be raise 1.4% lower by implementing the planner's allocation. Aggregate output would be raised 2.1%: one-third of the gains to output are negated by the increased congestion felt by residents of productive locations.

CHAPTER 2

The Impact of Emerging Climate Risk on Urban Real Estate Price Dynamics

with Matthew Kahn

2.1 Introduction

In the absence of a global agreement on reducing world greenhouse gas emissions, climate change risk continues to be exacerbated by ongoing population and per-capita income growth. Rising greenhouse gas production increases atmospheric concentrations of carbon dioxide and this increases the probability of extreme climate events. Researchers working at the intersection of macro and environmental economics have evaluated the ex-ante social costs of "fat tail" disaster events (Barro, 2013; Costello et al., 2010; Pindyck, 2013; Pindyck and Wang, 2013; Weitzman, 2009, 2011). Macro models of climate change's consequences have often implicitly assumed away any spatial adaptation possibilities.[1]

Recent research on climate change adaptation has studied how changing climate conditions affects regional comparative advantage (Desmet and Rossi-Hansberg, 2013a; Costinot et al., 2012). The former present a general equilibrium model featuring free

[1]Early work introducing a multi-region economy such as the Nordhaus and Yang (1996) RICE model did not allow for migration of labor across borders. In their model, regions differ with respect to their sectoral shares (i.e some regions have an agricultural focus while others specialize in manufacturing). They assume that the damage function from climate change is identical for each industry across different regions, and then conduct a shift-share calculation to determine how a region is affected by climate change. For example, if a region's economy specializes in agriculture and if climate change impacts agriculture then this region will be sharply impacted by climate change.

trade across regions and study the welfare consequences of climate change when the location of both production and households can shift as climate conditions shifts while the latter analyze the likely macro effects of micro-scale reallocations of cropland in the face of a changing climate.

Empirical studies have estimated cross-city real estate hedonic pricing regressions to predict how climate change is likely to impact the value of real estate in different locations (Albouy et al., 2013; Kahn, 2009). Intuitively, in the year 2050 will the current real estate price differential between San Francisco and Detroit narrow if Detroit's climate is predicted to improve (i.e. warmer winters) relative to San Francisco?

In this paper, we analyze the spatial implications of emerging climate risk within a system of cities, each of which may face different risks. Within a nation such as the United States or a trading bloc such as the European Union, both labor and capital can move to any location. This potential for spatial arbitrage imposes cross-restrictions across real estate prices and local wages across space. In spatial equilibrium, both heterogeneous households and firms cannot raise their utility and profits by moving to another location. Cross-city real estate prices and wages adjust to support such that the local land market and the local labor market clear (Rosen, 2002).

We introduce a system of cities model to consider the economic incidence of an emerging new catastrophic risk (i.e., climate change). Throughout this paper, we assume that a known subset of cities face the most severe risk due to their coastal geography. As climate scientists make progress and reveal the "new news" that cities such as Miami, New Orleans and New York City face increased risk of large-scale disasters, we seek to understand how equilibrium real estate prices across cities evolve. We embed the future risk of climate change into a classic Rosen/Roback (Rosen, 1979; Roback, 1982; Rosen, 2002) compensating differentials model. Climate change risk is a tied local public bad that (in expected value) poses future costs.

We contrast two main cases. In the first case, the population has homogenous preferences over coastal amenity attributes, climate, and avoiding low probability risks. The hedonic pricing gradient for homes immediately reflects the "new news" of the increased risk that cities such as Miami face due to rising greenhouse gas emissions.

In the second case, we introduce population heterogeneity along several dimensions. The population can differ with respect to income, tastes for amenities, the ability to engage in self protection against emerging risks (see Ehrlich and Becker, 1972), and location-specific networks and knowledge. Together these factors create a wedge between the willingness to pay among coastal incumbent home owners to remain in risky places versus the willingness to pay of outsiders considering moving to at risk cities. In this case, home prices in affected cities may not decline when the "new news" about climate risk becomes common knowledge. This result can be derived even when everyone agrees about the serious risk that the coastal cities face. An econometrician conducting an event study is likely to underestimate the average person's willingness to pay to avoid climate risk.

The key intuition here is to recognize that the marginal household, whose willingness to pay to live in the risky city sets the market price, may have a comparative advantage in coping with local risk or may have built up city specific capital (both social capital and local knowledge) such that this household effectively faces a higher migration cost for leaving the city. As we discuss below, this endogenous differential valuation of the same city by "insiders" and "outsiders" has implications for considering the merits of place based disaster insurance such as government FEMA programs.

In the literature on the fat tails of rare disasters, households within the model are aware of risks of which the econometrician is ignorant. In our model, the situation is different: both households and the econometrician are fully aware of the objective risks facing different cities. In this paper, the econometrician is unaware of the type and degree of household heterogeneity. This limits the extent to which prices of real estate—determined by marginal agents—reflect the willingness to pay of the average household. This limitation, due to residential sorting based on observed and unobserved attributes of both the city and the migrant, bears a similarity to the work of Shogren and Crocker (1991) on the attributes incorporated into hedonic pricing functions. Our findings concerning the information embedded in the hedonic gradient build on recent work by Kuminoff and Pope (ming) in studying the economic incidence of changes in local public goods (in this case coastal safety).

Our findings build on past work in local public finance by scholars such as Starrett (1981), who examined the conditions under which local public goods will be capitalized completely, partially, or not at all. In that context, Lind (1973) and Kanemoto (1988) provide guidance on the interpretation of capitalization studies as providing bounds on the heterogeneous population's willingness to pay for location specific attributes.

In the last section of the paper, we discuss how essential heterogeneity (see Heckman et al., 2006) affects inferences from standard hedonic real estate event studies where the event in question is the realization that specific cities face severe climate change risk. Our findings have a similar flavor as the Shogren and Stamland (2002, 2005) analysis of hedonic wage regressions seeking to recover statistical value of life estimates in the presence of essential heterogeneity.

2.2 Will Miami Vanish?

The motivating example for this paper is Miami. The Miami metropolitan area is home to six million people. The city itself is located six feet above sea level. In summer 2013, Rolling Stone magazine published a long front page article focusing on the claim that Miami is doomed because of imminent sea level rise (Goodell, 2013). This salient case study highlights the coming challenge that the U.S coastal population faces. Rappaport and Sachs (2003) document that a majority of the nation's population and income is located in coastal and Great Lakes areas. In the case of Miami, urban planning documents highlight that Miami-Dade County is planning for sea level rise (Miami-Dade County, 2010). The housing crisis notwithstanding, Miami home prices have increased nearly as rapidly as those of far-inland Denver over the last thirty years, showing no stark decline as climate research has progressed.[2]

The apparent non-responsiveness of Miami real estate to changing climate risk poses a puzzle: why aren't holders of Miami real estate assets compensated for this risk with a price discount? This puzzle is almost the inverse of the equity premium puzzle, where risk-averse investors hold bonds despite the low returns. The answer of Barro (2006) is that the fat-tails of consumption disasters, unobserved to the econometrician, lead investors to hold safer assets. In our puzzle, people pay apparently large sums to hold risky assets—Miami real estate—whose risk profile has increased with the advent of climate change. In our case, unobserved household heterogeneity means that only the households most willing or capable of dealing with these risks choose to hold Miami real estate, limiting the price impacts of emerging climate risks.

In considering the Miami case as a leading motivating example, we seek to focus attention to the damage natural disasters pose to the place-based capital stock rather than to human longevity. Cross-country research has documented that natural disasters are killing fewer people over time and that richer nations suffer fewer deaths from natural disasters (Kahn, 2005). We recognize that extreme natural disasters such as Hurricane Katrina which is estimated to have killed roughly 1,850 people in 2005 can be deadly. Valuing each life lost at $6 million yields a total value of life lost at $11.1 billion. Estimates for the property damage from Hurricane Katrina are in the range of $100 billion (Knabb et al., 2005). An alternative way to look at the damage caused by Katrina is to recall that in the year 2000 that New Orleans had a population of 490,000. This means that the 1,850 deaths from Katrina represented 0.0038 of the area's total population; for comparison there were 210 homicides in New Orleans in the same year (Van Landingham, 2007).

[2]See Figure 1 in the appendix. Data from Trulia indicates that Miami coastal areas such as Coral Gables and Miami Beach experienced an even more pronounced recent boom.

In the case of Hurricane Sandy in 2012, this storm caused 117 deaths and a total of $65 billion dollars of damage (Mulvihill, 2013; Newman, 2012). This example highlights that the damage risk to physical property swamps the total death risk. We believe that the ratio of total value of lives lost to climate change disasters divided by the total damage to physical capital will only decline over time. With the rise of smart phones and emergency warnings, we predict that the footloose coastal population (facing mandatory evacuations) will become more responsive to disaster alerts so that fewer people die in disaster events while buildings and infrastructure are highly immobile and subject to extreme damage. This discussion motivates some of the modeling assumptions we make below.

2.3 A Model of Rare Disasters with Variable Risk Across Cities

2.3.1 The Model with Homogeneous Households

Consider a model of household location choice where households maximize lifetime utility. To choose a location $j(t) \in J$ requires ownership of an asset h_j (i.e., a home) that provides access to city j's amenity a_j and also its idiosyncratic maintenance shocks. The first maintenance shock is a small but regular depreciation shock δ_j and the second is a rare but large catastrophic shock κ_j.[3] Both are i.i.d. across time, and independent across space.[4]

Households are free to buy and sell these assets each period, and transactions occur prior to the realization of any shocks. For a household currently living in city i who chooses to live in city j, lifetime utility is the discounted sum of period utilities, given

[3]In addition to housing capital, public capital is also at risk, and κ_j can be interpreted as including the risk that local homeowners will be compelled to rebuild damaged public capital. We discuss private capital in a later section.

[4]The same climatic pressures may create drought in one area and flooding in another, and a hurricane may impact multiple locales; the i.i.d assumption is a simplification. However, given that households choose one city at a time and they choose this city prior to the observation of shocks, their concern is with the average expected shock in their choice city, rather than possible correlations across cities.

by $u(c_{ij}, a_j)$. The houshold faces a period-by-period budget constraint of the form $c_{ij} + (p_j - \delta_j - \kappa_j)\, h_j = y + p_i h_i$ where y is the household endowment and p_i is the equilibrium price in city i. The subscript ij denotes a household that begins in city i and chooses to live in city j, recognizing that consumption may be different for two households who both move to j but who began in different cities.

For each city, there is a fixed supply of homes. This assumption can be interpreted as each city having a fixed quantity of land that must be combined with housing materials in a fixed proportion in order to produce housing services, and that the depreciation and catastrophic shocks represent regular maintenance and disaster-rebuilding, respectively. Alternatively, the assets can be interpreted as trees that bear two kinds of fruit: a fixed a_j and a variable $y - \delta_{jt} - \kappa_{jt}$. Seen in this light, the model is similar to those of Barro (2006, 2013), where we have allowed for a variety of asset trees from which households must choose just one.

In equilibrium, households will not choose to move; they will have already sorted into the city that maximizes their utility given relative prices. By assumption, each period's location decisions are made before the shocks are observed and thus relative prices are invariant from period to period and depend only on *expected* shocks.

We now consider the following asset pricing exercise to calculate the willingness to pay for real estate in different cities. If a household residing in city i purchases a home in city j for price p_{ij}, they expect to receive utility $U_{ij} = E[u(y + (p_i - p_{ij}) - \delta_j - \kappa_j, a_j)] + \sum_{t=1}^{\infty} \beta^t E[u(y - \delta_{jt} - \kappa_{jt}, a_j)]$. By setting $U_{ij} = U_i$, where U_i is the utility the household would receive if they stay, we can solve for the maximum price p_{ij} that the household would pay to move to city j.

$$\begin{aligned} E[u(y + (p_i - p_{ij}) - \delta_j - \kappa_j, a_j)] + \sum_{t=1}^{\infty} \beta^t E[u(y - \delta_{jt} - \kappa_{jt}, a_j)] \\ = \sum_{t=0}^{\infty} \beta^t E[u(y - \delta_{it} - \kappa_{it}, a_i)] \end{aligned} \tag{2.1}$$

In the initial scenario, the κ and δ processes are stationary, and the equation can thus be simplified.

$$\begin{aligned}(1-\beta)E[u(y+(p_i-p_{ij})-\delta_j-\kappa_j,a_j)]+\beta E[u(y-\delta_{jt}-\kappa_{jt},a_j)]\\ =E[u(y-\delta_{it}-\kappa_{it},a_i)]\end{aligned} \tag{2.2}$$

We suppose that at some initial date zero, households were free to choose locations and prices adjusted to make them indifferent. Those households who initially chose high-amenity and low-risk areas would have bid more initially and those high prices would persist in the steady state.

2.3.2 The Model with Heterogeneous Households

A similar equation can be derived in the case of heterogeneous households. While the potential types of household heterogeneity are limitless, we focus on three key cases: variable income levels, variable self-protection abilities, and the formation of local endogeneous social networks.[5] The first takes the form of a different endowment y_h, where h indexes households; the second involves a utility parameter ρ_h that reduces the effects of the catastrophic shock κ; and the third is modeled as a fixed moving cost μ. Solving the budget constraint for c and substituting again gives an equation that can be solved for the willingness to pay of household h in city i considering a move to city j.

$$\begin{aligned}(1-\beta)E[u(y_h+(p_i-p_{ij})-\delta_j-(1-\rho_h)\kappa_j,a_j)]+\\ \beta E[u(y_h-\delta_{jt}-(1-\rho)\kappa_{jt},a_j)]-(1-\beta)\mu=E[u(y_h-\delta_{it}-(1-\rho)\kappa_{it},a_i)]\end{aligned} \tag{2.3}$$

We can use this equation to solve for initial distribution of households across cities. We further assume that the distributions of $E[\delta_i]$, $E[\kappa_i]$, and a_i are such that *a priori*,

[5] The third case is reminiscent of Krupka (2009), where each household invests in human capital that enhances its particular local amenities.

all residents agree on which is the"worst" city, which we denote city l. Note that in equilibrium, $p_l = 0$. Consider then Equation (2.3) for household h in city l considering a move to city j. To solve for the distribution of households, consider each city $j \in J$ and calculate the willingness to pay p_{lj} for each household. Ordering these bids from highest to lowest, and recalling that each city has a fixed supply of homes, call $\hat{p}_j$ the willingness to pay of the marginal household. Take the city with the highest $\hat{p}_j$, and allocate to that city those households willing to pay at least $\hat{p}_j$. After allocating these households, repeat the process for the remaining households and cities until all households are allocated. This is the initial equilibrium distribution of households. Note also that the order in which the cities are selected offers an implicit ranking of the quality of life in these cities.

We use this distribution to solve for the set of prices across all cities. Beginning with the last two cities allocated. Because the worst city l has $p_l = 0$, the price of the next-to-worst city is the price that makes the marginal household in city l just indifferent between choosing the next city. Repeat this process for each city, moving up the implicit ranking identified previously. The last city priced will be the first city that households were allocated to: that with the highest marginal willingness to pay.[6]

After allocating households and solving for each city's prices, the resultant equilibrium will initially be stable: the relative values of the various shocks are stable across cities, and no household will be willing to pay to move a better city nor interested in paying less for less amenable city. This finding will not, in general, hold true in the next section after introducing climate change.

[6]Note that the equilibrium price p_j will in general not be equal to that calculated when allocating households across cities, $\hat{p}_j$. The former is calculated based on the marginal household's willingness to pay to move from their next-best city while the latter was calculated based on the willingness to pay to move from the worst city.

2.4 Climate Change Risk as "New News"

We now introduce climate change. Climate change is a one-time unanticipated event that alters the future risks of different cities in the economy. There is no learning *per se*: agents simply wake up to a new probability distribution of future outcomes. As an example, they may discover that climate change will affect Miami in the year 2040, Chicago in 2060, and Denver in 2080. They will also uncover the magnitude of the effect in each city. In particular, households learn that the distribution of catastrophic shocks $\kappa_{i,t}$ will worsen in the future for each city, as a function of the stock of global greenhouse gases. For each city, we suppose that there is a threshold level of greenhouse gases ϕ_i that will trigger the one-time transition from relatively-low to relatively-high risk, where the relative changes may vary by city. If greenhouse gases rise predictably, then this translates to a threshold year which we call τ_i at which point city i will increase in risk. We proceed with these assumptions in place.

2.4.1 Real Estate Pricing Impacts of Climate Change with Homogenous Households

The economy is in steady-state equilibrium when climate change is discovered. Once all cities have transitioned from their low- to high-risk state, the economy will be in a new steady-state equilibrium and a new version of Equation (2.2) will hold. We now consider what will happen to these bid functions during the transition to this new equilibrium. For convenience, we suppose that utility is quasilinear in the consumption good.[7]

With homogeneous households and fixed housing supply, prices in each period will ensure that no household will choose to move in any future period. With this equilibrium condition, we can write down the bid function for a household living in city i considering

[7]This simplification departs from Barro (2013), where the strict concavity of utility is critical in generating a premium for risky assets. Were we to maintain strict concavity, the null finding of no price change would be an even more surprising result.

a move to city j as a one-time choice between the discounted stream of amenities and shocks in city i and those in city j. The expected depreciation shock in city i is written δ_i and the expected catastrophic shock in city i is written κ_i^L or κ_i^H for low- and high-risk periods, respectively.

$$\begin{aligned} p_{ij} = p_i &+ \frac{1}{1-\beta}\left[(\delta_i - \delta_j) + u(a_j) - u(a_i)\right] \\ &+ \sum_{t=0}^{\tau_i} \beta^t \kappa_i^L + \frac{\beta^{\tau_i}}{1-\beta}\kappa_i^H - \sum_{t=0}^{\tau_j} \beta^t \kappa_j^L - \frac{\beta^{\tau_j}}{1-\beta}\kappa_j^H \end{aligned} \tag{2.4}$$

Willingness to pay to move from city i to city j is increasing with the relative amenity value in j and decreasing with the relative expected losses in j. Willingness to pay is decreasing in the transition year τ_i but increasing in the transition year τ_j. Note that cross-effects are not important in the homogeneous case as the price in any other city j' will be such that no household would be strictly better off by moving there.

It is clear from Equation (2.4) that any future changes in local climate will be reflected in house prices immediately. To an observer, the only difficulty in ascertaining whether the relative price has fallen in city j would be if τ_j were sufficiently distant that the discounted effects of climate change are negligible.

2.4.2 Climate Change's Impact on the Cross-City Spatial Equilibrium in the Essential Heterogeneity Case

We now investigate how our three types of household heterogeneity affect the equilibrium housing price dynamics in response to new information about the severity of climate change. We are especially interested in cases where a "neutrality result" holds, in which real estate prices of high-amenity but at-risk locales like Miami remain unchanged despite the discovery of climate changes that adversely affects such cities. As before, the three dimensions of heterogeneity are captured by household income y_h, household self-protection ability ρ_h, and the moving cost μ. The self-protection parameter takes a value between zero and one and measures the portion of a catastrophic shock that affects a particular household; a high value indicates that a household is not greatly affected.

Proposition 6. *Changes in relative climate risk across cities will alter the relative prices of those cities only if the changes in risk alter the willingness to pay of the marginal resident. If the marginal resident of an at-risk city possesses perfect self-protection capabilities, costly local endogenous networks, or a high enough income, then their willingness to pay will be unchanged and prices will not change despite an inarguable increase in climate risk.*

Income heterogeneity will produce *ex ante* sorting whereby the rich locate in (and bid up the price of) high-amenity cities. So long as the rich are rich enough they will choose to remain in high-amenity Miami after the new news of climate change—despite its increased risk of catastrophic shocks.[8] Because the choice set of cities is bounded, there exists a highest-amenity city. Suppose that utility is separable and strictly concave in consumption, and that for high-income households we have that $\frac{\partial u(c_i,a_i)}{\partial a_i} > \frac{\partial u(c_i,a_i)}{\partial c_i}$ for even high-amenity cities. In order to enjoy the best amenities in the country, the very wealthiest are willing to rebuild their houses every year.[9]

For those at the other end of the income scale, however, climate shocks will compound their already-high marginal utility of consumption. If the poorest are unable to bid their way ouf of low-amenity but high-risk locales, then they will suffer particularly large costs from climate change. Because of the limited ability to pay of the poor, and the already low amenity value of these locations, the fall in observed prices in these locales will be smaller than that observed in the case of homogeneous households—and smaller than the

[8]If the rich face meaningful risk of death, they will retreat to less-risky cities and leave the poor to enjoy the amenities in riskier cities. Avoiding risky cities is a type of input in the health and safety production function, and in this case the rich will place a greater value on avoiding risk (Hall and Jones, 2007).

[9]Of course, the best amenities might be in a low-risk city. If the unconditional distribution of disaster risk across cities is the same as its distribution conditional on amenity values, then rich will choose high-amenity but low-risk areas. It is the correlation of amenities (beaches) with risks (hurricanes) that leads to an underestimation of willingness to pay to avoid climate risk.

average household's willingness to pay to avoid climate change. Another way of saying this is that the middle class don't have to place large bids in order to outbid the poor for houses in safer cities. Income heterogeneity at both the upper and lower ends will thus serve to underestimate the average costs of climate change.

Returning to the quasi linear utility case, the possibility of heterogeneity means that the bid function for an individual of type h in city i for a home in city j must be modified:

$$p_{ij} = p_i + u(a_j) - u(a_i) - (E[\delta_j] - E[\delta_i]) - (1 - \rho_h)(E[\kappa_j] - E[\kappa_i]) - \mu \tag{2.5}$$

Self-protection against the risk of climate change provides a source for the neutrality result. In the extreme, a Miami resident with the ability to perfectly self-protect will exhibit no change in their willingness to pay for living in a risky city so long as its amenities are unaffected. Even in the face of seemingly extreme climate catastrophes, a Miami filled with such households will retain its initial price. A Denverite with no self-protection abilities will reduce her bid for Miami real estate one-for-one with the change in expected losses from catastrophic shocks. And indeed, they—like the econometrician—will be surprised to see that Miami residents with high self-protection show no inclination to leave nor to pay less to remain in Miami.

Finally, endogenous localized social capital provides a third source for the neutrality result. The presence of the moving cost—our stand-in for the endogenous formation city-specific social capital—induces a wedge between the willingness to pay of the marginal resident of Miami and the marginal non-resident who settles in an alternative locale. Before settling on cities, the two marginal households—one just within the margin and one just outside—have nearly identical bids for Miami property. The winning bidder values Miami at price p_m and so the household that just misses out values it at $p_m - \Delta$. Upon settling into Miami and its next best alternative, respectively, the marginal Miami resident and non-resident see their bids drift apart: due to the cost of moving, the non-resident would now bid only $p_m - \Delta - \mu$ while the resident would rather pay $p_m + \Delta$ than move to their next-best alternative.

The moving cost μ could also be interpreted as the cost of locating in any except the household's "preferred" city, where this preference is exogenously determined. For instance, some residents of Miami prefer it due to its proximity to other nearby countries, a plausible interpretation of the sizable populations of Cubans, Colombians, and Venezuelans that live in the area. A similar wedge would open in this case, and could justify the continued *increase* in population that Miami has seen despite the discovery of climate change.

Whether due to endogenous networks or exogenous preference, the moving cost produces a wedge between the bid functions of the marginal resident and and the marginal non-resident. This wedge between residents and non-residents implies that, so long as the increase in the (future, uncertain) costs of climate catastrophes are smaller than the (immediate, guaranteed) costs of moving and establishing a new social network, Miami residents may rationally choose to remain rather than to move.

2.5 Three Extensions

In this section, we sketch three extensions of the model that merit future research. In our basic model, a fixed supply of land meant that the only observable outcome variable was relative prices, and an endowment economy meant that the only actors were households. These extensions relax these strict assumptions and explore the consequences, while maintaining the core intuition that sorting by heterogeneous agents might limit the observed responses of real estate prices, wages, and migration to climate shocks.

First, consider the case of introducing a national government that engages in coastal maintenance, provides public goods and reimburses homeowners for some portion of catastrophic losses. To simplify our analysis, we have abstracted away from introducing governmental social insurance, such as FEMA, to protect at-risk cities using federal tax revenue.[10] At first glance such spatial subsidies create a spatial moral hazard effect as the federal government is implicitly subsidizing risk taking by those who love coastal locations.[11] In our model, there is a subset of the population who inelastically demand to continue to live in at-risk cities. This willingness to pay is due to idiosyncratic matching of households' preferences to the attributes of different locations and due to the endogenous networks built up over time, which the household knows it will lose if it moves away. In this sense, Miami solves a co-ordination problem: despite the fact that the city faces new risk its total package of attributes compensates for the risk and keeps the rational household in place. In such a setting, the benevolent government will recognize that those who remain are more "victims" than opportunists.

Within the model, the discovery of the catastrophic shock process comes as truly "new news": it's a zero-probability event against which agents cannot have insured themselves. Once discovered, agents are free to move elsewhere to avoid future climate change—if they can find a willing trading partner, a possibility that can only arise with population heterogeneity. For Miami, climate change will not induce in-migration as the city is now a worse prospect than before; at the same time infra-marginal residents may not wish to leave and, in any case, will find few willing partners in a sale. These facts suggest a role for a national government to invest in place-specific subsidies—whether defensive protections like sea walls or transfers in the event of a catastrophe—for Miami and, so long as the subsidy is not too great, there is no concern of moral hazard.

A second modification to the model would be the introduction of endogenous local housing supply. Our formal model focuses on the housing demand side and simply fixes an inelastic housing supply, which implies that any change in the marginal willingness to pay

[10] Popular Flood Insurance Law Is Target of Both Political Parties

[11] Kousky et al. (2006) discuss the interaction between government place based investments and household locational choice. In their model, multiple equilibria emerge as the government is more likely to build seawalls if more people are expected to live there and more people will move to an area where sea walls are expected to be built.

would cause an immediate change in prices. This price sensitivity gives substance to our neutrality results. However, endogeneity of housing supply also enables the possibility of net population changes in high-risk and low-risk locations—an additional source of data to the researcher.

Even with endogenous housing supply, the durability of housing capital will nevertheless yield a kinked housing supply curve as presented in Glaeser and Gyourko (2005). They argue that in a city such as Detroit the durability of housing means that there is a fixed supply of existing older homes (built when Detroit was the car capital of the world) yielding a vertical supply curve up to the point where the price of housing exceeds the marginal cost to developers of building new housing. In this endogenous-housing extension to our model, any Miami resident who prefers to leave following the discovery of the catastrophic shock process would be able to do so. As in the Detroit of the Glaeser and Gyourko (2005) model, durable housing capital will remain in place despite these evacuees: supply is downwardly inelastic as before, and thus as before prices in Miami are sensitive to any decrease in willingness to pay.

A final extension of the model would be to introduce local labor markets in which firms hire workers, rent land, and invest in city-specific capital. Firms might face different self-protection costs than households, and might expect different reimbursement from government programs. However, firms would also sort spatially: service-sector firms with low capital requirements (like households with effective self-protection) may face a negligible penalty from locating in Miami. This sectoral sorting would tend to keep wages high in Miami despite climate risks to Miami's capital goods. Furthermore, high-amenity but risky locales may specialize in attracting wealthy retirees who earn capital income from safer regions. This will create a small but well-compensated labor pool in high-amenity at-risk areas. In all of these cases, the inclusion of firms generates additional dimensions along which sorting acts to minimize the observed changes in at-risk locales. Conversely, capital-intensive firms may follow their workers to at-risk locales. In this case, the cost of insuring their capital will lower the wage offers in risky areas.

2.6 Empirical Implications for Hedonic Research Measuring Disaster Capitalization

Economists often estimate dynamic hedonic models to test if real estate prices change in response to changes in local public goods such as air quality improvements (Chay and Greenstone, 2005), Superfund site cleanups (Greenstone and Gallagher, 2008; Gamper-Rabindran and Timmins, 2011), and improvements in urban transport infrastructure (Zheng and Kahn, 2013). Our model of emerging climate risk, where the discovery of this risk is new news to households, lends itself to an event study framework. We now seek to position our paper's findings within this existing literature.

Consider the following hedonic regression, where ϕ_i is a continuous variable that indexes susceptibility to climate risk. In particular, ϕ_i is the threshold level of CO_2 that will trigger a transition from the low- to the high-risk state in city i. For example, suppose that $\phi_{Miami} = 450ppm$. When the level of atmospheric CO_2 reaches 450 ppm, the risk and severity of climate disaster will undergo a one-time increase from their current levels. For simplicity, suppose that all cities face identical initial climate risk, that ϕ_i triggers an identical increase in disaster risk for all cities, and that the interest rate is constant.[12] The only variation across cities comes from the timing of the transition, which is governed by ϕ_i; a "risky" city is therefore one that will experience climate change sooner than later.

Suppose the econometrician observes sales prices for a large sample of homes with the same physical structure scattered across a range of cities over many year. The econometrician observes each city's quality of life attributes and each city's susceptibility to climate disaster as indexed by ϕ_i. However, the econometrician does not observe

[12] Within our model, these simplifications amount to a common κ for all cities before climate change, a new (but still common) κ after, and the previous assumption of quasilinear utility.

household characteristics of those buying and selling houses.[13] Finally, we denote by τ the year of climate change discovery, and $\mathbf{1}_\tau(t)$ takes the value 1 only for the year τ. Under these assumptions, we can write down an event study regression model, where the change in price upon the discovery of climate change is regressed upon the index of climate risk suspectibility.

$$Price_{h,i,t} - Price_{h,i,t-1} = a \times Z_i + b \times \mathbf{1}_\tau(t) \times \phi_i + U_{h,i,t} \tag{2.6}$$

In this regression, b represents the compensating differential for a higher threshold for climate change. In the context of our model with homogeneous households, the average person in the economy would be willing to pay b to avoid the extra maintenance costs and the marginal increase in the death risk associated with a lower threshold of ϕ_i—that is, associated with additional time spent under the high-risk climate regime. The regression coefficient b should reflect the payment that keeps her just indifferent in expected lifetime utility.

Now consider the case in which people differ with respect to their incomes, their self-protection capabilities, and their localized social capital. Suppose that there are two types of cities, coastal (e.g., Miami) and inland (Denver) where the threshold ϕ_i is smaller in coastal cities. The price of coastal real estate is fixed by the willingness to pay of the marginal coastal resident. In the extreme case of self-protection in which the marginal coastal resident can perfectly offset climate disasters, the marginal bid for coastal real estate would not change at all upon the discovery of climate change and the researcher would thus recover an estimate of zero for b and conclude that markets are not pricing risk. This is the empirical counterpart to the neutrality result given in Proposition 6.

Conversely, if the marginal Miami resident has no self-protection abilities, then the price of Miami real estate will fall and the researcher will conclude that Miami residents

[13] We are assuming that households are buying and selling homes (perhaps because of life cycle considerations) and this generates the sales data that the econometrician observes.

are suffering large climate-triggered losses—even if the typical Miami resident is a self-protector who faces limited utility costs from climate change. In this case, the unobserved variation in the ability to self protect against catastrophic risks will create the appearance that coastal households are exposing themselves to a high degree of risk, relative to the price discount they receive for this exposure. These results have a similar logic as that of Shogren and Stamland (2002, 2005) who focused on what can be inferred from conventional value of a statistical life hedonic wage regressions in the presence of population essential heterogeneity (Heckman et al., 2006).

Populations may also differ in unobserved location-specific demand. This possibility further complicates the interpretation of the hedonic real estate regression presented above. Households that build valuable city-specific social networks may lose access to this social capital if they leave, and even the marginal coastal household may have a discontinuous willingness to pay for their current city relative to the alternatives. When climate change is discovered, the increased climatic shocks will represent an expected cost to these households and yet the marginal coastal household may choose to remain if the economic rents exceed the climatic losses.[14] In this case, the dependent variable—the price change of real estate—does not provide any insight into the underlying demand curve, nor the changing welfare of the residents of at-risk cities in light of climate change.

A similar case emerges when income heterogeneity leads the extremely wealthy to sort into Miami due to its high amenity value. After accounting for the increased costs of climate catastrophes, the very rich nevertheless have a higher marginal utility from amenities than from consumption. In the high-income limit, climate change therefore has no effect on their bid functions for Miami real estate. As in the case of endogenous social networks, the fact that prices do not change after the advent of climate change does not necessarily imply that underlying welfare is unchanged.

[14]Due to their status as port cities and the historical (and ongoing) roles as entry points for immigrants, many coastal cities feature large ethnic enclaves that generate valuable social capital for major population segments. This suggests that coastal cities, differentially susceptible to climate change, may also have populations with differentially strong social ties.

These examples show that inferences to be drawn from observations of risk capitalization are limited, but the regressions are not useless: the bid of the marginal resident places bounds on costs of climate change for both residents and non-residents. The marginal willingness to pay to avoid the risks of coastal cities can thus be interpreted as an upper bound for the willingness to pay for a typical resident, and a lower-bound for the typical non-resident.

Consider again the case with two city types: coastal and inland. The initial change in prices upon the discovery of climate change will produce upper and lower bounds for the willingness to pay to avoid risk for coastal and inland residents, respectively. After discovery, the prices of both cities will decline over time as the onset of climate change nears due to the dwindling number of "low-risk" periods; these price changes could tighten the bounds on willingness to pay. The rates of price change in the two city types may change again after coastal cities transition to high risk,[15] providing additional information, and the eventual relative prices after all cities have transitioned to high-risk may further illuminate the scope of these bounds.

2.7 Conclusion

Climate change is likely to pose different costs on different cities. Coastal cities and cities located close to rivers will face greater flood risk while other cities such as Phoenix may face extreme summer heat. Such dynamics in location specific attributes suggests that forward-looking asset markets such as real estate should reflect the present discounted value of these relative risks.

In this paper, we have introduced an equilibrium system of cities model in which households hold common expectations of spatial variation in the risks that different cities face. We document that a standard event study research design will yield very

[15] For instance, due to high-income individuals evacuating high-amenity coastal cities only after they transition to high-risk.

different estimates of the risk premium for being exposed to extra climate change risk depending on the degree of household heterogeneity. While in standard asset pricing, asset risk contains no idiosyncratic component and the CAPM style model captures the risk premium, in the case of housing—one's home bundles both an asset's rate of return and one's access to a specific city's attributes and to the social connections one has built in that location. This idiosyncratic match (either on unobserved tastes or endogenously built up social capital) creates a wedge between how an insider values remaining in the area versus how others in the society value the asset (Miami) now that the new news about climate change is common knowledge. We document that owners of Miami real estate are now faced with abnormally high risks, but—unlike in the case of risky equity—they do not appear to receive a large compensation for bearing this risk.

The model has implications for event study style hedonic real estate research. In the presence of the three dimensions of heterogeneity that we have presented, an empirical researcher's reduced form estimate of risk capitalization will provide bounds on the willingness to pay for avoiding new risks (Bajari and Benkard, 2005), and further changes in relative prices may narrow these bounds. Our findings highlight the key role of explicitly modeling the residential sorting process (Kuminoff et al., 2013).

CHAPTER 3

Entrepreneurship, Information, and Growth

with Stephan Weiler, Eric Thompson, and Sammy Zahran

3.1 Introduction

In this paper, we explicitly quantify and incorporate entrepreneurship into a spatial-equilibrium endogenous growth framework to better understand the role of entrepreneurs in determining economic growth. We emphasize the revelation of marketplace information and a resultant externality, aspects of entrepreneurship that have not been identified previously in the literature. Through the successes and failures of their projects, entrepreneurs generate valuable marketplace information regarding the contours of the geographic and industrial territory in which their projects reside. This externally beneficial information can be utilized by future entrepreneurs, who can emulate successful projects and avoid the pitfalls identified by the failures.

We quantify this understanding by focusing not on the entrepreneur but on her project. To undertake a project, an entrepreneur must assess local and broader demand for her products and services, the necessary supply network, and the feasible financing. Each entrepreneurial action—opening, expanding, or even terminating a project—illuminates a niche of the marketplace. Project financiers must assess these same aspects as well as the suitability of the entrepreneur herself. Through its evolution and possible eventual demise, the project itself provides information on the viability of similar projects.

Our basic hypothesis is maintained empirically. Spatial differences in entrepreneurial activity, as expressed through the births of local establishments, have a statistically and economically significant effect on local employment growth. This finding is robust to a variety of specifications. Furthermore, local establishment births *and deaths* are strongly associated with future establishment birth rates—evidence that future entrepreneurs look to past successes *and failures* when choosing what projects to implement. Finally, we close the loop by identifying twin indirect positive impacts that establishment deaths have on future employment growth within metropolitan counties—effectively underscoring the likely informational role of entrepreneurial projects on growth in these populous counties. We find similar results when we include establishment births alongside existing measures of entrepreneurship; these results support our core contention that establishment births capture a relationship with employment growth that is different from its relationship to alternatives such as proprietorship and small business employment.

The fundamental relationship between entrepreneurial projects and subsequent employment growth remains consistent when using instrumental variable analysis featuring historical mining activity as a first-stage instrument. As in Glaeser et al. (2012), we follow the Chinitz (1961) hypothesis that large-scale mining precipitated a large-scale industrial structure inimical to entrepreneurship. These results are robust to a variety of specifications of independent variables and remain remarkably strong even after including traditional controls known to influence local growth.

We proceed with a discussion of the existing literature and the motivation for studying entrepreneurial projects within an endogenous growth framework, and present the model. Then, we test the intermediate hypothesis that levels of entrepreneurship are positively dependent on past establishment births and deaths before proceeding to the main growth analysis. The results are highly supportive of our main hypothesis that entrepreneurship causally supports economic growth, and further evidence suggests that the marketplace information framework is useful for interpreting these findings. The paper closes with a discussion of these results and a brief conclusion.

3.2 Motivation and Background

Recent studies like Audretsch and Keilbach (2004, 2007) have provided substantive evidence of a relationship between entrepreneurship and growth, while Ács et al. (2009) and Braunerhjelm et al. (2010) provide models featuring entrepreneurs who implement innovations flowing from spilled-over knowledge and research. These approaches update earlier endogenous growth models including those of Krugman (1987, 1991), Lucas (1988), and Romer (1986, 1990) wherein knowledge or human capital are created endogenously and external effects of differential knowledge cause divergent outcomes across economies. Ács and Armington (2006) emphasize entrepreneurial processes at the local level, echoing the hypotheses of Lucas (1988) and Jacobs (1970). Our paper corroborates these findings and provides evidence that revealed marketplace information is an important mechanism in accounting for the relationship between entrepreneurship and growth.

Information has a distinctly local characteristic in such a formulation (e.g. Weiler, 2006). The potential for business success is market-specific, both in its potential demand contours (e.g. for non-traded service industries) as well as more universally in its supply/cost character (e.g. labor, land, capital, insurance, among others). In regions with "thinner" markets—featuring fewer openings, closings, and other business transactions—information on the potential of such markets will consequently be more limited. While such information gaps may themselves impede prospective entrepreneurs directly, they may also indirectly restrict business opportunity through higher perceived uncertainty by critical loan and insurance suppliers. In the spirit of Akerlof (1970), we thus see differences in market-specific entrepreneurial experience as driving *geographic information asymmetries*, which can in turn yield suboptimal local investment and growth.

Previous work incorporates revealed marketplace information within a geographic context. Lang and Nakamura (1993) apply the geographic asymmetry logic to neighborhood housing markets where numerous transactions produce precise information, reduce future-period uncertainty, and encourage future activity. Weiler (2000a) provides a game-theoretic model and case study of information revelation where, upon the success of a pioneering firm, project viability is revealed and other firms make entry decisions. Weiler et al. (2006) extend the theoretical structure of entrepreneurial information revelation in a Bayesian Revealed Preference framework, where the perceived probability distribution of project outcomes is updated through discrete information increments.

The logic of geographic information asymmetries is borne out in the financial literature. Distance is a significant factor in determining abnormally high returns in firm acquisitions, with more localized transactions allowing greater insight into the most promising targets (Basu and Chevrier, 2011). These findings are particularly strong for relatively small, non-public, R&D-intensive, non-metropolitan firms with no analyst coverage—precisely the most informationally opaque to non-locals (Uysal et al., 2008). These high returns notwithstanding, Weiler et al. (2006) find that the social benefits of information provision are nevertheless likely to exceed the private benefits, suggesting a potential market failure.

Local context presents an opportunity for the study of economic growth amongst regions with broadly similar institutions and populations. International macroeconomists view integrated local economies as potentially revealing contexts for better understanding the micro-economic mechanisms of the macroeconomic growth process (e.g. Krugman, 1991), while labor economists are leveraging local labor market analysis in pursuit of this same lens (e.g. Moretti, 2011). Human capital and its localized external effects have been emphasized by Shapiro (2006), Moretti (2004), Glaeser et al. (1995) and Hammond and Thompson (2008). Such work is further motivated by the tremendous variation in economic performance of subnational economies, a question which has challenged the economics field for decades (e.g. Hall et al., 1970; Summers et al., 1986; Blanchard et al., 1992; Weiler, 2000b). Low and Weiler (2012) in fact suggest that entrepreneurs take these variable conditions into account when choosing to become self-employed, as their decisions are directly influenced by the risk and returns of wage and salary employment opportunities in local labor markets.

Entrepreneurship may be one of the fundamental microeconomic mechanisms leading to geographically asymmetric outcomes—indeed, similar reduced-form analysis by Stephens and Partridge (2011) and Stephens et al. (2013) suggests that the lagging Appalachian region could best be boosted by a focus on entrepreneurship. We hold that the finding is general in that it applies nationally, and provide evidence that entrepreneurial activity and marketplace information specifically are critical to economic growth. The next section presents a model of spatial equilibrium and endogenous growth to aid in assessing whether entrepreneurial decisions to undertake, expand, and terminate microeconomic projects determine the path of local economic development.

3.3 Model and Data

We outline below a model of endogenous growth that is premised on the Roback (1982) model of a spatial equilibrium resulting from the maximizing choices of households and firms. Following Stephens et al. (2013), we consider labor migration flows and their determinants. For households, the key choices revolve around relative real wages and local amenities, while relative wages and productivity are determinative for firms. Structural factors combine with agents' decisions to produce equilibrium wages, house prices, and ultimately net labor flows; these flows become the object of our reduced-form analysis.

Households maximize utility U by choosing from a range of cities, indexed by i, the location that maximizes U based on the expected wages for the household's relevant skills, the prices of non-traded land and services, place-based amenities, and other local and idiosyncratic characteristics. Household choices of optimal locations will produce a long-run equilibrium in which the utility flows from locating in different regions are equalized for identical households. In the medium run, changes in local labor supply will depend positively on relative local utility ($U_i - U$, where U is their utility elsewhere) net of moving costs (M_i) as formalized in (3.1), where $L^S(.)$ is increasing in its argument.

$$\Delta Labor\ Supply_i = L^S(U_i - U - M_i) \tag{3.1}$$

Firms maximize profits π by choosing from a range of cities the location to produce that maximizes π based on an array of considerations: the local wage, the size and skill composition of the labor market, the availability and cost of local non-traded inputs (e.g. land), the area's access to markets both internal and distant, and other local productivity determinants including agglomeration economies. When choosing a location, firms weigh these factors and locate where expected profit is greatest. In the long run, entry and relocation will equilibrate expected profits; in the medium run changes in local labor demand will depend positively on relative profits ($\pi_i - \pi$, where π is their profit elsewhere) as summarized in (3.2), where $L^D(.)$ is increasing in its argument.

$$\Delta Labor\ Demand_i = L^D(\pi_i - \pi) \tag{3.2}$$

Equations 3.1 and 3.2 can be combined to create a reduced form relationship between growth in the local labor market and the underlying determinants of household-side amenities and firm-side productivity effects. In the reduced form, local employment growth is a function of local economic characteristics that influence the growth of labor supply or demand. Characteristics that influence utility will affect growth in labor supply. Natural amenities influence utility, as does proximity to a metropolitan area given the variety of job openings and social and cultural opportunities afforded by cities. Likewise, local characteristics that influence profitability will affect growth in labor demand. Demand shocks related to the industries present in a region influence profitability as do characteristics that determine productivity such as the human capital of the local workforce and agglomeration economies resulting from the scale and scope of local economic activity.

Equation (3.3) shows the reduced-form relationship between local employment growth and its various determinants. Following the above discussion, these determinants include household-side amenities (A_i), measures of human capital including worker education (BA_i) and occupation profile (OC_i), idiosyncratic productivity effects or demand shocks (Z_i)—and others including the share of individuals in creative occupations (CC_i) and market access (MA_i).[1] To the list of employment growth determinants we append the propensity to engage in entrepreneurial projects (EP_i).

$$Growth_i = G(A_i, BA_i, OC_i, Z_i, CC_i, MA_i, EP_i) \tag{3.3}$$

Entrepreneurial projects are included in Equation (3.3) in accord with our hypothesis that entrepreneurship will increase productivity of the local economy through information spillovers. Any entrepreneurial project will likely involve management and resource allocation, innovation, financial risk-taking, and other tasks and chores.[2] Rather than emphasize any particular aspect of the entrepreneurial process, the experimentation involved in a project is the core of our theoretical understanding of entrepreneurship—whether that experimentation involves a new product or process, a new allocation of resources within an industry or region, or any other change from the previous economic arrangements of a locale. Such a project has the potential to create social benefits beyond its private benefits in the form of increased marketplace information. Through their success and failures, entrepreneurial projects reveal the shape of market demand, supply

[1] Physical capital—along with wages and other prices—is excluded here as maximizing agents will equilibrate its return across locales and so it cannot play a role in the reduced-form empirical implementation.

[2] Left untouched is the source of inspiration for entrepreneurial projects. Knowledge spillovers from research and development are one possibility, but we remain agnostic on the ultimate source of project inspiration.

conditions, and the viability of concepts or innovations from elsewhere to work within a particular context. They enable established firms, future entrepreneurs, and financiers to make better-informed decisions, whether they choose to replicate or support successful projects or to improve upon or avoid the pitfalls highlighted by unsuccessful ones. This improved information directly affects the productivity of enterprises within a locale.

Based on this project-oriented informational perspective, we identify establishment births and deaths as our key measure of entrepreneurship. An establishment birth aligns nicely with the theoretical understanding of an entrepreneurial project: both brand-new firms and new locations for expanding firms involve the experimentation highlighted above, and both induce the informational externalities at the core of this understanding of entrepreneurship. Establishment deaths generate information about types of entrepreneurial projects a local economy could not support. These metrics present an improvement on traditional measures such as the share of small firms or of the self-employed, both of which include many stagnant enterprises while neglecting the information content provided by establishment births and deaths.

As suggested by the reduced form growth model, we utilize the empirical strategy shown in Equation (3.4) below:

$$\begin{aligned} Growth_i = {} & \beta_0 + \beta_1 * Entrepreneurial\ Projects_i + \beta_2 * Other\ Entrepreneurial \\ & + \beta_3 * Amenities_i + \beta_4 * Share\ with\ BA_i + \beta_5 * Share\ High\ Human\ Capital_i \\ & + \beta_6 * Demand\ Shock_i + \beta_7 * Age_i + \beta_8 * Density_i + \beta_9 * Income_i \\ & + \beta_1 0 * Employment + \beta_{11} * Distance\ to\ Metro_i \\ & + \beta_{12} * Lagged\ Employment\ Growth + \beta_{13} * Lagged\ Population\ Growth + \varepsilon_i \end{aligned} \tag{3.4}$$

The control variables follow our reduced form model and existing literature including

Stephens and Partridge (2011); Stephens et al. (2013) and also Hammond and Thompson (2008), Rappaport (2004, 2007), Partridge et al. (2009), Beeson et al. (2001), Ács and Armington (2006), Huang et al. (2002), Deller et al. (2001) and Glaeser et al. (1995). The variables include education attainment, the share of employment in high human capital and creative occupations, natural amenities, population density, initial income and employment, previous population and employment growth, median age, proximity to metropolitan areas, and industry mix employment growth—the last a proxy for the local demand shock derived from the combination of initial county-level industry mix with national-level industry trends.

Hammond and Thompson (2008) and Glaeser et al. (1995) emphasize the contribution of education attainment to subsequent county and metropolitan area growth. We utilize a measure of education attainment based on obtaining a bachelor's degree (or higher) as well as two measures of human capital based on occupation: the share of workers in high human capital occupations and in occupations in the arts. The classifications follow the USDA's Economic Research Service definitions of creative class occupations and its artistic subset, respectively.[3] We treat the two as separate shares, so that the artistic occupations—those classified as "art and design workers" as well as "entertainers ... and related workers"—are excluded when calculating the high human capital share of employment.

Rappaport (2004, 2007) and Beeson et al. (2001) suggest persistent population growth in regions with higher amenities, which is supportive of employment growth. The amenity variable is taken from the USDA's Economic Research Service and reflects January and July temperatures, humidity, sunshine, topography and water coverage. Variables reflecting proximity to metropolitan areas were developed following Stephens and Partridge (2011). These variables reflect the propensity of counties to benefit from the spillover of urbanization economies generated by metropolitan areas of different sizes. The growth

[3] We prefer "high human capital" to "creative class" as the latter seems to imply a Bohemian character while the actual data includes accountants, managers, and lawyers.

benefits of all distance variables are expected to decline with distance.

Partridge et al. (2009) include an industry mix employment growth variable as a measure of the local demand shock, a key control. For each industry j, we calculate the national growth rate ($G^j_{2000--2007}$) for the period 2000–2007, and for each county i, we calculate the initial employment share in each industry (s_{ij}). The demand shock ($D^i_{2000--2007}$) is the sum of the products of national growth and initial local share: $D^i_{2000--2007} = \sum_j G^j_{2000--2007} * s^{ji}_{2000}$. Industry employment totals are at the three-digit NAICS level, except for farming and government employment, which are taken at the broad sector level. For suppressed values, we implement a simplified version of the technique outlined in Isserman and Westervelt (2006). Given initial employment, this local demand shock reflects the employment that would follow had each local industry behaved in line with its national counterpart. The variable will thus capture any trends related to the prominence of industries at the local level that are experiencing nation-wide growth or decline. Our entrepreneurship variables are left to explain the residual of local growth after controlling for the intersection of national trends with local industrial composition.

Our use of establishments birth and death data for entrepreneurial projects in Equation (3.4) warrants a brief discussion. We draw the establishment data from the Business Information Tracking Series. The birth and death rates are reported in units of births (or deaths) per thousand employees for the period 1998–1999, which is the earliest available period. In some regressions, we include the product of local establishment births and deaths to enable the measurement of dual effects from establishment deaths: the direct negative closure effect and the indirect positive information effect. When paired with the birth rate, the simple death rate includes both the negative closure effect and the positive information effect. Alone, the product summarizes the state of local entrepreneurial dynamism by effectively synthesizing overall entrepreneurial activity that generates useable market information; a high product and considerable information flow is possible even with a small net birth rate. When the birth rate, the death rate, and their product are combined, we are able to interpret the two opposing impacts of the death rate individually, with the death rate capturing the closure effect while the product isolates the information effect. A positive regression coefficient for the product variable provides a channel by which an additional death may have a measurably positive information effect capable of partially or even wholly offsetting the inarguable negative closure effect.[4]

While we view the novel entrepreneurial variables as an improved metric, their novelty argues for external confirmation. Three more common measures of entrepreneurship are thus included as well: the share of proprietors in total employment, the ten-year growth in the count of proprietors weighted by initial employment, and the estimated share of employment in firms with four or fewer employees. We include these measures for contiguity with other research and as a check of our hypotheses while noting potential drawbacks: proprietorship is a legal definition and the choice to operate as a corporation may reflect tax policies rather than intrinsic entrepreneurship, while a business's small size may simply reflect a lack of desire or ability to expand. We include these variables first alone, then each individually alongside the establishment birth rate, and finally altogether with the birth rate.

To alleviate concerns that we may misinterpret a non-causal relationship, we use an instrumental variable approach. Our instrumental approach utilizes historical mining employment data taken from the 1974 County Business Patterns. Our key instrument is the log of the estimated employment in the mining industry. Although mining employment shares may seem more appealing as instruments to reveal structures antithetical to entrepreneurship, the discrete existence of a non-trivial mining sector in fact may be more important in framing critical local market transactions. In particular, the higher wages and land rents inherent in such heavily-embedded mining operations create barriers to entry in the form of fundamental benchmarks for local factor markets, shifting the

[4]There is a high degree of collinearity between their product and births and deaths themselves. To account for this in the regression analysis, the interaction is the product of de-meaned births and deaths. The interpretation of marginal effects remains the same.

decision locus of both workers and firms towards such opportunities but away from the development of an entrepreneurial dynamic (Weiler, 1997, 2000b, 2001). Potential entrepreneurs are simultaneously burdened by these high costs while receiving few benefits from the large scale (and largely exported) mining activity (Graves et al., 2009). Within the spatial equilibrium model outlined previously, mining raises wages and land rents for all firms but raises productivity (and lowers input costs) only for those industries that either utilize the mines or directly benefit from their output—for instance, steel manufacturing.

To this end, we also use two non-linear instruments based on mining employment. First, we deploy indicator variables for counties with mining sector employment greater than different thresholds, which enables identification of areas where mining activity is likely to be export-oriented and thus pressure local factor prices. Second, we leverage the interaction between log mining employment and population density. While each of these instruments produces similar results when used individually, their use together produces the clearest perspective and so we focus our attention on the multiple-instrument regressions rather than on the single-instrument regressions, or on alternatives like the share of employment in the mining industry.[5] Finally, the period 1974 is chosen as it both has readily-available mining data and is sufficiently distant to be plausibly exogenous to employment growth in the post-2000 period, making it a potentially valid instrument.

One concern with the instrument may be that it proxies for locations in structural decline for non-entrepreneurial reasons. To counter this possibility, we bolster the case for the validity of the instrumental approach by including the local demand shock and past employment growth as controls. In its construction, the first variable—which is the industry-mix growth term from shift-share analysis—directly utilizes 1998 local industry employment at the six-digit NAICS level, including mining industries. Therefore the regressions are analyzing solely the 2000–2007 "regional shift" from shift-share analysis.

[5] To accommodate multiple potentially endogenous variables, in some instances we also utilize deep lags of alternative measures of entrepreneurship based on local proprietorship data.

Furthermore, our inclusion of 1990–2000 employment growth as a control means that we are effectively testing for a relationship between the portion of the regional shift that is uncorrelated with local employment growth in the prior decade, further guarding against concerns that the instrument. If the lagged mining instruments were solely proxying for structural decline, the industry mix and lagged employment growth variables should render the instrument impotent. By accounting for industry structure in 1998, the first-stage relationship between our mining instruments and the residual local growth is more likely to be capturing the posited relationship

Due to the inclusion of these stringent controls—and in some regressions, the inclusion of state fixed effects—we see few alternative explanations for a relationship between historical mining activity and current employment growth besides our explanation: that historical mining activity dampened entrepreneurialism. As we shall show in the results section, multiple statistical tests confirm this hypothesis.

The main analysis focuses on employment growth from 2000-2007. The initial year is chosen for two reasons: first, it enables the use of control data from the 2000 decennial census, and second, it enables a peak-to-peak analysis across the business cycle. Non-census control variables are also from 2000. Before proceeding to these regressions, we explore the relationship between establishment births and deaths in 2000 and births in 2005. Our information hypothesis predicts that both births and deaths should be positively associated with future births. These regressions include the same control variables as the main regressions; the industry-mix growth term is adjusted to capture the period 2000–2005.

The primary unit of analysis is the county; in many instances we focus on the subsets of metropolitan or non-metropolitan counties. Counties offer the smallest-scale check of the hypotheses, given the lack of data available at smaller units. As counties vary in size from fewer than 100 residents to nearly ten million, we use initial-period employment to weight the observations for both the summary statistics and the regressions. We obtain

Dependent Variables	Mean	Std Dev	Min	Max
Employment Growth Rate, 2000-2007	8.75	11.7	-35.3	169
Employment Growth Rate, 2000-2011	6.28	13.5	-38.2	214
Entrepreneurship Variables	Mean	Std Dev	Min	Max
Establishment Births, 1998-99	4.32	1.1	0	12.6
Establishment Deaths, 1998-99	3.95	0.9	0	17.2
Establishment Births, 1998-99 (Metro Counties)	4.34	1.04	0.49	10.4
Establishment Deaths, 1998-99 (Metro Counties)	3.93	0.84	0.43	9.08
Establishment Births, 2005	4.72	1.33	0	23.1
Establishment Deaths, 2005	4.05	1.06	0	28.9
Log of Mining Employment (plus one), 1974	5.2	2.27	0	9.98
Mining Employment Greater than 20, 1974	0.84	0.36	0	1
Mining Employment Greater than 100, 1974	0.62	0.48	0	1
Population Density, 1980	2.4	8.39	0	62.2
Share of Proprietors, 1998	15.4	4.81	1.74	56.9
Share of Proprietors, 1979	12.2	4.01	0.78	41.1
Employment-Weighted Proprietor Growth, 1988-98	4.62	4.77	-13.7	99.3
Employment-Weighted Proprietor Growth, 1969-79	6.07	6.65	-8.17	106
Employment Share in Small Establishments, 1998	5.15	2.06	0.65	100
Employment Share in Small Establishments, 1974	4.25	2.08	0	100
Control Variables	Mean	Std Dev	Min	Max
High Human Capital Share, 2000	24.6	7.14	0	53
Arts Share, 2000	1.25	0.81	0	6.67
Bachelor's Degree Share, 2000	26	9.85	4.92	60.5
Tract-Weighted Population Density, 2000	6.5	15.4	0	114
Log Income, 2000	10.6	0.28	8.03	11.4
Log Employment, 2000	12.4	1.68	4.55	15.5
Employment Growth, 1990-2000	22.2	20.6	-39.4	767
Predicted Employment Growth, 2000-2007	6.5	2.82	-16.5	86.4
Median Age, 2000	35.2	3.2	20.6	54.3
Population Growth, 1950-1960	0.39	0.48	-0.42	3.71
Amenity Score	1.3	3.35	-6.4	11.2
Distance to Nearest MSA in miles	26	31.8	0	371
Marginal Distance to MSA > 250,000	41.7	64.3	0	762
Marginal Distance to MSA > 500,000	58.3	81.5	0	797
Marginal Distance to MSA > 1,000,000	78.3	97.8	0	797

Table 3.1: Summary Statistics. Observations are weighted by year 2000 employment unless otherwise noted. Observations total 3,072 for all variables except the metropolitan county subset for which there are 825 observations. See the data appendix for more details.

similar results when using metropolitan areas as the unit of analysis; these results are reported in Appendix B.

The summary statistics for all variables are presented in Table 1. Additional detail about the data and variables used can be found in Appendix A.

3.4 Entrepreneurship and Information

We hypothesize that entrepreneurial projects create socially beneficial information about marketplace opportunities and thus induce faster economic growth. Through the failures and successes of entrepreneurial projects, local observers glean information that they are then able to put to use in their own enterprises and funding decisions. In this sense, entrepreneurship breeds more entrepreneurship, promoting economic growth.

We first directly test our hypothesis by exploring whether measures of entrepreneurial projects have a positive link to future establishment births. For this analysis, we utilize a similar framework to that presented above, and estimate a nearly identical regression to that implied by (3.3). In place of employment growth through 2007, the dependent variable is the establishment birth rate between 2005 and 2006.[6] Alongside the identical suite of control variables, we utilize the measures of entrepreneurial projects discussed above: the establishment birth rate, death rate, and their product. We expect to find a positive effect from all three variables as entrepreneurs and financiers are more willing to engage in risky entrepreneurial projects when those projects are made less uncertain by the information yielded from the success and failure of prior projects.

In Table 3.4, we present results that suggest this is the case. The first column suggests that a unit change in the birth rate yields a response in the latter-period birth rate of

[6]Similar results hold whether 2005 establishment birth rates are taken as total births divided by 1998, 2000, or 2005 employment. The alternatives ensure that the result is not mechanical; these results are available upon request.

Birth Rate, 1998	0.87***	0.7***	0.68***	0.62***
	(0.025)	(0.038)	(0.036)	(0.09)
Death Rate, 1998	—	0.23***	0.23***	0.35***
		(0.039)	(0.04)	(0.1)
Births × Deaths, 1998	—	0.036**	0.01	0.038
		(0.015)	(0.014)	(0.026)
N	3072	3072	3072	825
R^2	0.78	0.79	0.82	0.94
F-test for non-FEs variables	272	242	90	27.5
Counties	All	All	All	Metro
Fixed Effects	No	No	State	MSA
Full controls	Yes	Yes	Yes	Yes
Regression	OLS	OLS	OLS	OLS

Table 3.2: Preliminary Entrepreneurship Results. Dependent variable is the establishment birth rate, 2005. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 3.1 and the data appendix for details on the variables. See Table B.2 in the appendix for full results including for control variables.

over 0.87—a very high degree of correlation, even in the presence of additional controls.[7] Alongside the establishment birth rate, we introduce the establishment death rate their product in the second column. Both the prior-period establishment death rate and the product of births and deaths have effects that are positive, statistically significant, and economically sizable: a one-point increase in the initial death rate is associated with an increase in the final birth rate of greater than one-fifth. For the interaction, the significance in the results of Column Two suggest that the magnitude of churning, whereby new enterprises replace or subsume older ones, has an independent positive effect independent of net establishment growth in the initial period. The birth and death rates remain significant in the presence of state fixed effects (Column Three) for all counties, and even for MSA fixed effects (Column Four) when examining the metropolitan subset. The significance of any churning effects on future births apparently become incorporated into state and MSA fixed effects, as indicated by the final two columns. Full results for

[7]Complete results with control variable coefficients are available in the appendix for all regressions.

the model are available in Table B.2.

3.5 Mining as an Instrumental Variable

For the employment growth regressions to follow, concerns about forward-looking entrepreneurs suggest the possibility of endogeneity. In the following section, we utilize historical mining activity as an instrumental variable. The logic, as described above, is that large concentrations of mining employment were likely to induce an economic ecosystem inimical to entrepreneurial activity. Mining activity will make for a useful instrument as, after including our stringent array of control variables, historical mining employment is likely to have little effect on employment growth—except through the entrepreneurship channel. However, its very strength as an instrument in the employment regressions makes the measure ill-suited as an instrument in the previous regressions. If (a lack of) mining activity induces a self-sustaining equilibrium with (high) low rates of entrepreneurial activity, then that mining activity cannot be thought of as exogenous to the dependent variable of future-period entrepreneurship. And, indeed, when mining is included in the regression alongside the three measures of entrepreneurial activity—births, deaths, and their product—the coefficient on mining is negative and significant at the 5 percent level and thus not exogenous.[8]

This finding suggests that mining activity is very strongly related to our chosen measure of entrepreneurship, suggesting that it may be a useful instrument in the growth regressions. To this effect, the first-stage results are presented in Table B.2 in the appendix. These results, wherein we regress modern establishment birth rates on modern controls and the various instruments—including the key instruments of the log of 1974 mining employment—are strong and fit close to theory. In the main regression analyses, we use a suite of mining-related variables: the log of 1974 mining employment, nonlin-

[8]These results are available from the authors upon request.

ear indicator variables for employment greater than 20 and 100 workers, an interaction between log of 1974 mining employment and local population density in 1980, and population density itself in 1980. The last two instruments follow from the hypothesis that smaller, less dense locales will have less latent entrepreneurial activity to lose, and that mining activity in dense counties will thus have a stronger negative impact; these coefficients have the predicted signs and are statistically significant in most specifications. For some regressions, we also make use of lagged versions of the alternative entrepreneurship measures based on proprietorship and small business employment.

The theoretical and empirical strength of the establishment birth and death results suggests that the combination of our novel measure of entrepreneurial information and the historical mining IV strategy offers a useful approach to analyzing the impact of entrepreneurship on economic growth. Furthermore, the uniformly strong results, including the positive relationships found between prior-year deaths and future establishment births, provide compelling initial evidence that an information effect is at work, whereby past failures significantly inform and foster future business creation. We utilize these findings and interpret the employment growth results in this light.

3.6 Results

The rate at which entrepreneurial projects are undertaken as measured by establishment births, deaths and their product is hypothesized to have a positive and significant effect on local employment growth. This is precisely what we find. We highlight the main specifications in Table 3.6 and include a range of alternative specifications to establish the robustness of the findings in the various appendix tables. Throughout the various permutations, the core result stands: that entrepreneurial projects have a significantly positive effect on employment growth. Finally, we incorporate the establishment birth rate, the establishment death rate, and their interaction to show that deaths have a measurable positive effect beyond the direct negative impact (as an establishment closure directly lowers employment). These results are shown in Table 3.6.

The first two columns of Table 3.6 present OLS results, the first column without and the second column with state fixed effects. The last two columns show the same results while using instrumental variables, the third column without state fixed effects and the fourth column including them. The results are remarkably consistent and consistently strong: for example, using the first column results, a one standard deviation increase in the establishment birth rate is associated with a 4.0 percentage-point increase in the growth rate of employment. With a mean employment growth rate of 8.75 percent, this is a meaningful movement. Furthermore, the stability of the coefficient to the introduction of state fixed effects in the second column suggests that the control variables are capturing a substantial portion of the variation across places and that the estimates are, in fact, capturing a relationship between entrepreneurial projects and employment growth.

One might be concerned that forward-looking entrepreneurs simply anticipate future growth and that these results suffer from reverse causality. These concerns are lessened by the similarly strong results from the third and fourth columns, which show IV results. First, the mining instruments are strong. Using the Cragg-Donald Wald F-statistic to test the strength of the instrument, the statistics are 47.5 and 45 for the two specifications. Second, the Hansen J-statistic is not significant, suggesting that the instruments are themselves exogenous to the estimating equation.[9] Finally, the GMM distance test fails to reject the null hypothesis that the establishment birth rate is exogenous (with P-values 0.55 and 0.17, respectively).[10] We thus focus the discussion on the OLS results.

To test the hypothesis that establishment births capture a unique and novel compo-

[9]In Table B.2 in the appendix, we further test the exogeneity of the instruments by examining whether the inclusion of the establishment birth rate alongside the mining instruments reduces the (negative and significant) reduced-form relationship between the mining instruments and employment growth. In all cases, the mining instruments show no statistically significant relationship after the inclusion of the establishment birth rate—as expected.

[10]The GMM distance statistic tests whether the OLS and IV coefficient estimates are identical; rejection would have suggested that the IV estimates are appropriate.

Birth Rate, 1998	3.62***	3.61***	3.24***	2.18**
	(0.39)	(0.38)	(1.04)	(0.97)
Predicted employment growth, 2000–07	0.61***	0.55***	0.64***	0.63***
	(0.11)	(0.1)	(0.13)	(0.12)
Lagged Emp Growth, 1990-2000	0.2***	0.22***	0.21***	0.23***
	(0.038)	(0.045)	(0.041)	(0.047)
Log Employment, 2000	-0.27	-0.22	-0.26	-0.21
	(0.43)	(0.43)	(0.44)	(0.44)
Log Income, 2000	-12.2***	-14.1***	-12.9***	-16.8***
	(3.43)	(3.18)	(4.37)	(4.37)
Density, 2000	0.1***	0.03	0.1**	0.037**
	(0.035)	(0.029)	(0.04)	(0.031)
HC Share, 2000	0.75***	0.71***	0.78***	0.91***
	(0.28)	(0.24)	(0.28)	(0.27)
Arts Share, 2000	-1.8**	-0.12	1.81**	-0.039**
	(0.9)	(0.78)	(0.9)	(0.77)
BA Share, 2000	-0.37***	-0.4***	-0.38**	-0.377***
	(0.14)	(0.11)	(0.14)	(0.12)
Pop Growth, 1950-1960	0.61	0.53	0.72	0.79
	(0.83)	(0.72)	(0.91)	(0.76)
Median Age, 2000	-0.55***	-0.78***	-0.52***	-0.69***
	(0.11)	(0.1)	(0.13)	(0.12)
Amenity Score	0.11	-0.27	0.14	-0.1
	(0.14)	(0.24)	(0.17)	(0.26)
Distance to MSA	-0.017*	-0.004	-0.02	0.001
	(0.009)	(0.009)	(0.01)	(0.01)
Marg dist MSA > 250k	-0.004	-0.015***	-0.004	-0.015***
	(0.004)	(0.005)	(0.004)	(0.006)
Marg dist MSA > 500k	-0.004	-0.002	-0.004	-0.001
	(0.004)	(0.005)	(0.004)	(0.005)
Marg dist MSA > 1M	0.003	0.005	0.003	0.002
	(0.003)	(0.004)	(0.004	(0.005)
Constant	129***	158***	136***	193***
	(32.8)	(31.2)	(42.4)	(43.7)
N	3072	3072	3072	3072
R^2	0.55	0.64	0.55	0.63
F (non-FEs)	59.0	41.0	55.2	41.7
Counties	All	All	All	All
Fixed Effects	No	State	No	State
Full Controls	Yes	Yes	Yes	Yes
Regression	OLS	OLS	IV	IV
Weak IV-Robust F-test	—	—	47.5	45.0
Weak IV-Robust P-value			0.02	0.41
Endog of E'ship P-value	—	—	0.55	0.17
Endog of IVs P-value			0.25	0.99

Table 3.3: Main Results. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See Table 3.1 and the data appendix for details on the variables. The instruments used are as follows: the log of mining employment in 1974, indicator variables for mining employment greater than 20 and 100 in 1974, the population density in 1980, and the interaction between log of mining employment and population density. The weak instrument-robust P-value is for the Anderson-Wald F-test statistic; the weak instrument test is the Cragg-Donald Wald F-statistic. The test for the endogeneity of entrepreneurship (i.e., the birth rate) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J-statistic. See Table B.2 for robustness tests using alternative definitions of entrepreneurship alongside establishment births. Further robustness tests are available from the authors upon request.

nent of local entrepreneurialism, we have executed the same regressions as above while adding to the mix three traditional measures of entrepreneurship: the alternative measures are based on the share and growth rate of proprietorship and the share of employment in small businesses; these results can be found in Table B.2.[11] When including these measures as covariates, the birth rate results are virtually unchanged. Alone in OLS regressions, the alternatives show their expected positive signs, but the introduction of the establishment birth rate attenuates their effects. When using instruments for both the alternatives and for the establishment birth rate, no positive and significant relation- ship exists between any of the alternatives and employment growth. The birth rate, on the other hand, shows the expected positive relationship to employment growth in all cases. Entrepreneurial activity, as captured by the establishment birth rate, has a consistently strong positive effect on local economic growth.

As noted above, the covariates explain a significant portion of county-level variation in employment growth. By and large, these control variables show the expected relationships. For predicted employment growth, we find the expected positive and statistically significant relationship with observed employment growth in all cases. Workforce characteristics also significantly influence local job growth rates. In most cases, there is a negative relationship between the share of the workforce in arts occupations—entertainers and art and design workers—and employment growth.[12] A positive and statistically significant relationship is identified for high human capital occupations as classified by the USDA and including such occupations as managers, attorneys, and accountants. There is also an interesting interplay between the alternative methods of measuring human capital

[11] Results for further sensitivity tests are available upon request from the authors. These results consistently show a strong relationship between establishment births and employment growth with stable point estimates. In particular, the measure remains similar in magnitude even alongside the introduction of spatial dependence of both error terms and the dependent variable itself, as well as for the Census-designated metropolitan and nonmetropolitan subsets of counties. Nor does the period of measure for employment growth alter the results; using growth from 2000–2011 produces similar results.

[12] Intriguingly, Table B.2 reveals that the arts occupations' share of total employment may be positively related to the establishment birth rate—suggesting a possible relationship between so-called "creative classes" and entrepreneurial activity.

in counties: by occupation (as with HC Share) or by education (as with BA Share).

Employment growth was slower in counties with higher initial nominal incomes, consistent with the findings of Deller et al. (2001) for rural counties and Glaeser et al. (1995) for cities. No relationship was identified between initial employment levels and subsequent employment growth, but there was persistence in employment growth across time with positive and significant coefficient estimates on the 1990–2000 employment growth variable. When looking across the broad set of results included in the appendix, the census tract-weighted population density generally has a positive and statistically significant relationship, although this finding is reversed for non-metropolitan counties. Results also indicate slower employment growth in counties with a higher initial age. This result may reflect less potential for labor force growth among counties with an older population, although we note that Stephens and Partridge (2011) did not find a relationship between average age and employment in Appalachian and adjacent counties. Conditional on these other variables, no robust relationship was identified between amenities and employment growth.

One of the core predictions of the theoretical framework is that the termination of an entrepreneurial project provides information as to the local economic environment, and this information can be used by future agents in ways that should have a positive effect on economic growth. We have seen in the previous section that establishment deaths—our empirical approximation of project termination—do indeed have a positive effect on the future rate of entrepreneurial activity, and that, still further, the product of births and deaths itself has a positive effect. In Table 3.6, we explore whether the information derived from establishment deaths has measurable implications for subsequent employment growth.

Despite the handicap that establishment deaths have a direct negative effect on local employment, we find tentative evidence of the information-driven positive effect of establishment closures on eventual job growth, at least in metropolitan counties. The

Birth Rate, 1998	4.73***	5.2***	5.37***	7.74*	7.31*
	(0.56)	(0.75)	(0.75)	(4.32)	(4.35)
Death Rate, 1998	-1.84***	-3.2***	-2.97***	-9.44*	-6.11
	(0.5)	(0.84)	(0.84)	(5.07)	(4.64)
Births×deaths, 1998	0.17	0.71**	0.49**	1.99**	1.05
	(0.19)	(0.28)	(0.22)	(0.94)	(0.84)
N	3072	825	825	3072	3072
R^2	0.56	0.65	0.74	0.44	0.6
F	59.8	50.6	30.7	41.5	35.6
Counties	All	Metro	Metro	All	All
Fixed Effects	No	No	Yes	No	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes
Regression	OLS	OLS	OLS	IV	IV
Weak IV-Robust F-test	—	—	—	5.98	7.86
Weak IV-Robust P-value				0.046	0.13
Endogeneity of Entrepreneurship P-value	—	—	—	0.5	0.99
Endogeneity of IVs P-value				0.99	0.9

Table 3.4: Results with Birth and Death Rates. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See Table 3.1 and the data appendix for details on the variables. The instruments used are as follows: the log of mining employment in 1974, indicator variables for mining employment greater than 20 and 100 in 1974, the population density in 1980, and the interaction between log of mining employment and population density, the lagged proprietor's share of employment in 1979, the employment-weighted growth rate of proprietorship from 1969–1979, and the share of employment in small businesses in 1974. For first-stage results, see Table B2. The weak instrument-robust P-value is for the Anderson-Wald F-test statistic; the weak instrument test is the Cragg-Donald Wald F-statistic. The test for the endogeneity of entrepreneurship (i.e., the birth and death rates) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J-statistic.

product of births and deaths has a positive coefficient for the metropolitan county subset and when using instruments—although the insignificance of establishment births in the instrumental variable regressions clouds interpretation. In the instrumental variable regressions, the three potentially endogenous variables are jointly significant and, for the regression with state fixed effects, we can reject at the 10 percent level that the instruments are weak and we fail to reject that the entrepreneurship measures are exogenous. We therefore focus again on the OLS regressions.

For metropolitan counties, the positive interaction provides a channel whereby a marginal establishment death increases the employment growth rate. That the interaction's positive effect shows up most clearly in the subset of metropolitan counties is not surprising: large metropolitan counties, with their deep labor markets, provide insurance against idiosyncratic job losses. Employees at closing metropolitan establishments can more easily find a job, enabling the informational effects of such closures to be identified more readily.

Of course, the direct effect of those establishment deaths is negative; a closure is still a closure. But in this model, the marginal effect of an establishment death in a particular county is the sum of the death rate coefficient and the product of that county's establishment birth rate with the "B*D" coefficient. For metropolitan counties, this comes to $-3.2 + (B_i - 4.34) * 0.71$ where B_i is the local establishment birth rate and 4.34 is the mean of metropolitan county birth rates. This implies that for a metropolitan county with a birth rate above 8.8, the marginal death has a *net positive* effect on future employment growth. Despite the different coefficient estimates, the comparable figure for the IV point estimates is 10.4. Although the closure job loss effect isn't fully offset for the large majority of counties that are below this threshold, even in these counties each incremental establishment death carries with it a smaller negative employment bite.

This finding runs contrary to the intuition that knock-on effects from the loss of one establishment might have negative repercussions for others; the reverse is in fact the case.

Furthermore, as deaths are themselves highly correlated with births, all else is not equal: a marginal death carries with it a substantial fraction of an offsetting birth, further inducing employment growth. Along the lines of the theoretical framework outlined above, these *highly dynamic economies* feature an informational effect of establishment deaths that partially or fully offsets the negative direct effect.

The information content of deaths also helps to shape subsequent births, and the flip side of the noted result reinforces this implication of the information hypothesis. For the subset metropolitan of counties, the marginal effect of an establishment birth is estimated to be $5.2 + (D_i - 3.93) * 0.71$ where D_i is the establishment death rate and 3.93 is the mean death rate for metropolitan counties. A marginal birth thus has a larger effect on employment growth in the presence of more establishment deaths.

The information hypothesis again provides a framework for interpreting this result. Entrepreneurs and financiers in counties with many establishment deaths are able to draw on a deeper pool of failures as they shape new projects, sharpening their plans to avoid the pitfalls highlighted by previous failures and enhancing their projects' growth potential. The fact that new projects lead to more jobs in the presence of more establishment deaths underscores the significance of business failures to the growth process, a paradox which itself highlights both the role and importance of information flows.

While not dispositive, these twin novel findings of the indirectly positive impact of firm deaths fit comfortably within the marketplace information framework, dovetail with the results from the preliminary bridging results on the informational role of past entrepreneurship, and support the interpretation that entrepreneurial projects themselves have a causal effect on economic growth. The apparent propensity of entrepreneurial activity to sustain itself suggests that localities are subject to *geographic information asymmetries* that manifest both within a place's level of entrepreneurship but also, ultimately, in its longer-term level of economic development.

3.7 Conclusion

This paper tests the proposition that entrepreneurial projects have a positive and determinative effect on employment growth in local economies. We introduce a novel set of entrepreneurial measures tailored to the information hypothesis: establishment births, establishment deaths, and their product. Using an instrumental variable approach ensures that causation is well-established, and after controlling for a variety of structural factors, the results indicate that entrepreneurial projects are indeed a causal determinant of future growth. This result holds for both counties and metropolitan areas and while using different control variables, different time periods, state fixed effects, models with spatial dependence, and alternative measures of entrepreneurship. Consistent with the information hypothesis, the strength of the relationship between establishment births and employment growth in the presence of alternative measures suggests that the relationship captures a unique component of entrepreneurialism.

We also provide evidence that current establishment births feed off of nearby past entrepreneurial projects: the establishment birth and death rates, as well as their product, enter positively in determining future rates of establishment births—even in the presence of a full suite of control variables. These results are entirely consistent with the entrepreneurial information framework outlined in the paper. This framework, in which future entrepreneurs, financiers, and existing firms draw information from the successes and failures of entrepreneurial projects, provides a consistent lens to understand the surprising yet revealing positive results from establishment deaths. The latter finding suggests that future participants utilize the richer information set generated by these failures to strengthen and accelerate their own eventual entrepreneurial innovations. Regional variations in such information sets leads to *geographical informational asymmetries*, which themselves create reinforcing cycles of business (in)activity and consequent economic growth (stagnation).

If entrepreneurial projects sustain employment growth and are themselves partially sustained by the information generated by entrepreneurial failures, then those past failures ought to have a positive and measurable effect on employment growth. And, within metropolitan counties, firm deaths do have indirect positive effects on employment growth. These positive effects flow through two channels. First, information from a broader set of past failures will allow the marginal firm birth to have greater job-creation effects. Second, the most *highly dynamic economies* with sufficiently high establishment birth rates may see faster employment growth from the marginal establishment death, while establishment closures in less dynamic economies produce a positive informational effect that partially offsets the negative direct effects. These twin novel findings—alongside the instrumental variable and other analyses—provides powerful support to the notion that information from entrepreneurial projects is an important causal input to local economic growth.

These results underline the importance of entrepreneurship in economic development and push the limits of the current conception of entrepreneurship. Understanding entrepreneurship as a process of project-based information revelation—in addition to the well-known role it serves as a conduit for innovation—encourages both policymakers and researchers to see entrepreneurship anew, and to incorporate the unique and crucial actions of entrepreneurs into both academic research and practical policy.

APPENDIX A

Appendix to *Is The Rent Too High?*

A.1 Additional Figures

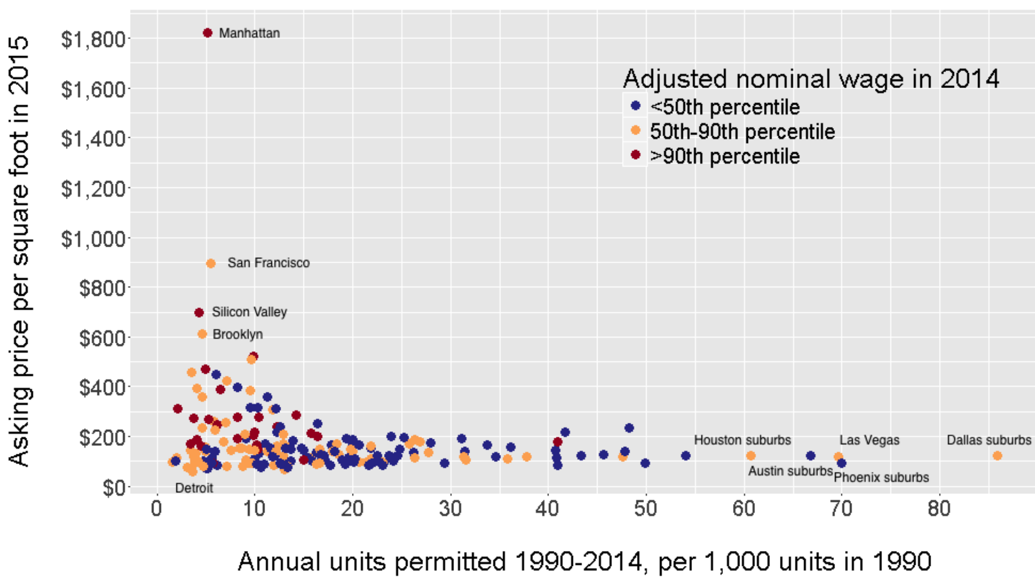

Figure A.1: Plot of housing prices and increases in the quantity of housing units in selected counties. In descending order, the top three counties are San Francisco, San Mateo, and Santa Clara: Silicon Valley.

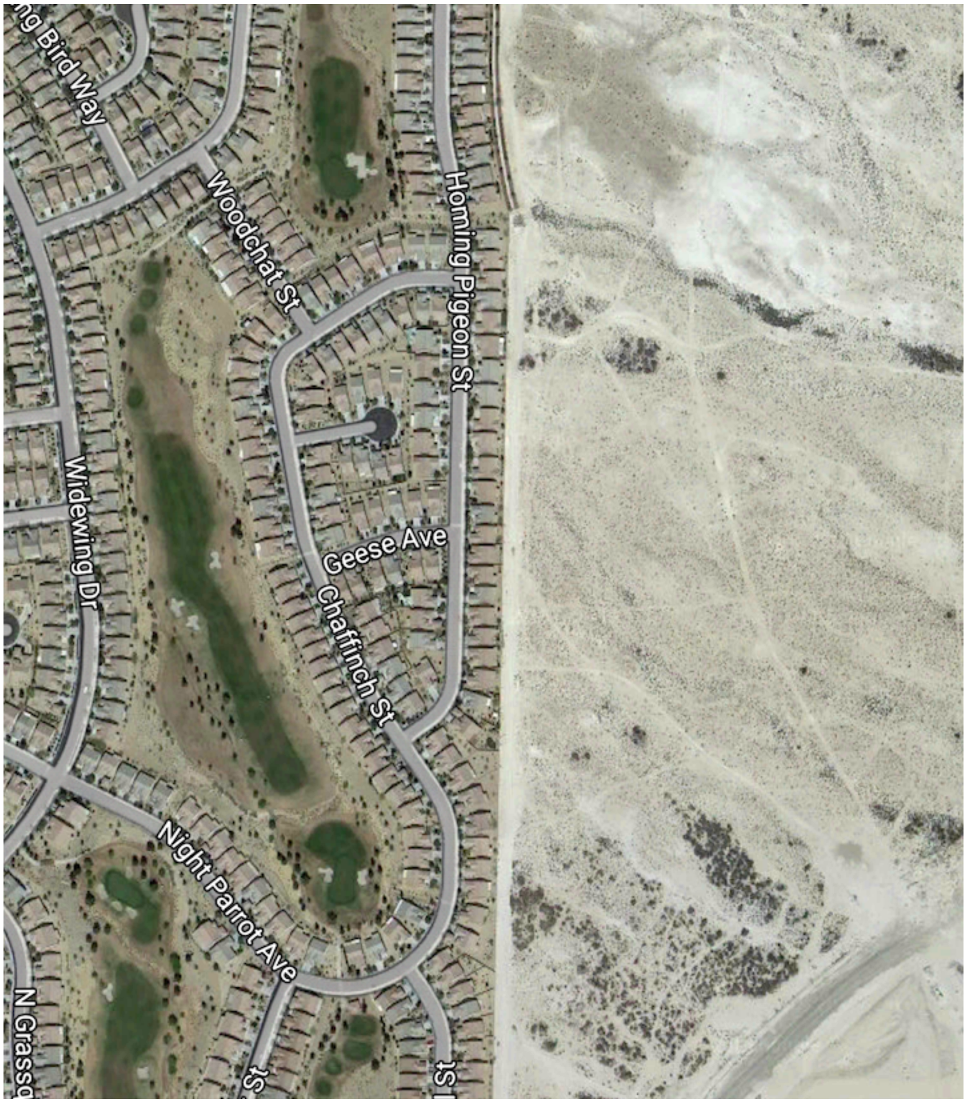

Figure A.2: A Google Earth satellite image of the built-up edge of Las Vegas. Note that the city edge is obvious: on the left, population densities are in the upper single-digit thousands while the right is undeveloped desert. While this desert is itself likely to fill in, the pattern of a sharp edge will likely remain. While desert development may entail unusually large fixed costs, similar patterns hold in regions with friendly climates.

A.2 Proofs

Proposition 1. *For a given x, the portion of the $OL(n,x)$ curve with $n > 0$ is strictly concave, hump-shaped, and intersects the $MC(n)$ curve twice, not at all, or once at a point of tangency.*

Proof. Consider an arbitrary $x > 0$. Recall that $OL((n,x)$ for $n > 0$ is defined as

$$OL(n,x) = y(x) + \Pi - \frac{F}{n} - u^{-1}\left(v(n) + \bar{u}\right).$$

As $F > 0$, $-\frac{F}{n}$ is strictly concave. As $u(c)$ is concave and strictly monotone, $u^{-1}(\cdot)$ is convex and strictly monotone. Similarly, $v(n) + \bar{u}$ is convex and strictly monotone for a given $\bar{u}$. Their composition is thus convex as well, implying that $-u^{-1}\left(v(n) + \bar{u}\right)$ is concave. The first two terms do not depend on n, and so $OL((n,x)$ is strictly concave in n.

By assumption, $v(0)$ is bounded below. Thus we have:

$$lim_{n\to 0} OL(n,x) = -\infty.$$

From above, $-u^{-1}\left(v(n) + \bar{u}\right)$ is concave and decreasing in n, so that

$$lim_{n\to \infty} OL(n,x) = -\infty.$$

Hence $OL(n,x)$ is hump-shaped.

Finally, consider the relation between $OL(n,x)$ and $MC(n)$. As marginal cost is linear in n, the hump shape of $OL(n,x)$ implies that for a given x they must intersect twice, once with tangency, or not at all. □

Proposition 2. *Consider an arbitrary x for which there exists n with $OL(n,x) \geq$*

$MC(n)$. For this x, the locally optimal zoning constraint solves:

$$\max_{\{c_0(x),\bar{n}(x)\}} u(c_0(x)) - v(\bar{n}(x)) \tag{A.1}$$

subject to the budget constraint

$$p_0(x) + c_0(x) + \frac{F}{\bar{n}(x)} = y(x) + \Pi \tag{A.2}$$

and the participation constraint

$$u\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - 2\bar{n}(x)/Z^2\right) - v\left(\bar{n}(x)\right) \geq \bar{u}. \tag{A.3}$$

Proof. Note that the participation constraint can be rewritten as

$$OL\left(\bar{n}(x), x\right) \geq MC\left(\bar{n}\right).$$

By Equation (1.6), satisfying the participation constraint implies $n^*(x) = \bar{n}(x)$: the zoning law chosen by the initial residents will become the equilibrium population. For $\bar{n}(x)$ not satisfying the participation, the equilibrium population will be $n^*(x) = 0$ if $\bar{n}(x)$ is to the left of the intersection of the $OL\left(n, x\right)$ and $MC(n)$ curves or $n^*(x) = n_H(x)$, the rightmost intersection of $OL\left(n, x)\right)$ and $MC(n)$, if $\bar{n}(x)$ is to the right. The first case cannot maximize the initial resident problem: due to the assumption that F is large, they will always prefer to have a positive population. In the second case, the eventual equilibrium population will be $n_H(x)$, which satisfies the participation constraint. That is, choosing $\bar{n}(x) > n_H(x)$ cannot improve upon simply choosing $\bar{n}(x) = n_H(x)$. Therefore there exists a zoning law $\bar{n}(x)$ that satisfies the participation constraint and that maximizes the underlying initial resident problem:

$$\max_{\{c_0(x),\bar{n}(x)\}} u\left(c_0(x)\right) - v\left(\theta\left(\bar{n}(x);\bar{u},\Pi\right)\right)$$

subject to

$$c_0(x) = y(x) + \Pi - \frac{F}{\theta\left(\bar{n}(x);\bar{u},\Pi\right)} - p_0(x)$$

It remains to see that the second-order conditions for a maximum are satisfied. Note that $u(\cdot)$ is concave and utility is separable in c and n. Writing as $\lambda_0(x)$ and $\mu_0(x)$ the Lagrange multipliers for the budget and participation constraints, the first order condition for $\bar{n}(x)$ is

$$\begin{aligned} & -v'\left(\bar{n}(x)\right) + \lambda_0(x)\frac{F}{\bar{n}(x)^2} && \text{(A.4)} \\ & + \mu_0(x)\left[u'\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - 2\bar{n}(x)/Z^2\right)\left(\frac{F}{\bar{n}(x)^2} - 2/Z^2\right) - v'\left(\bar{n}(x)\right)\right] = 0. \end{aligned}$$

As $v(\cdot)$ is convex, the second order condition for a maximum is satisfied as long as

$$\begin{aligned} & -v''\left(\bar{n}(x)\right) - 2\lambda_0(x)\frac{F}{\bar{n}(x)^3} + \mu_0(x)u''\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - 2\bar{n}(x)/Z^2\right)\frac{F}{\bar{n}(x)^2} && \text{(A.5)} \\ & - \mu_0(x)\left[u'\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - 2\bar{n}(x)/Z^2\right)\left(\frac{2F}{\bar{n}(x)^3} + 2/Z^2\right) + v''\left(\bar{n}(x)\right)\right] && \text{(A.6)} \\ & < -2/Z^2\mu_0(x)u''\left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - 2\bar{n}(x)/Z^2\right). && \text{(A.7)} \end{aligned}$$

Every term on the left hand side of Equation (A.7) is negative. This equation is easy to check for particular functional forms. For example, it is satisfied whenever

$$\frac{F}{\bar{n}(x)^2} > 2/Z^2.$$

□

Proposition 3. *The set of occupied locations* $\mathcal{X}$ *is* $[\underline{x}, 1]$ *for some threshold location* $\underline{x}$*. For the threshold location* $\underline{x}$*, the* $OL(n, \underline{x})$ *and* $MC(n)$ *curves are tangent at a unique level of population that depends on* $\bar{u}$*.*

Proof. Consider an occupied location $\tilde{x}$. As $y(x)$ is strictly increasing in x, $OL\,(n, x)$ must also be strictly increasing in x. For all locations $x > \tilde{x}$, this implies that there exists a positive n such that $OL\,(n, x) > MC(n)$. As initial residents prefer a strictly positive population, they will choose such an n and, by Equation (1), that location will be occupied.

Recall that by assumption, the productivity and fixed cost parameters are such that a *proper* subset of locations are occupied. By Equation (1.6) and Proposition 2, all locations with n such that $OL\,(n, x) \geq MC(n)$ will be occupied in equilibrium. By the continuity of $y(x)$, this set is closed. Thus there exists a threshold $\underline{x}$ such that $OL\,(n, \underline{x})$ and $MC(n)$ are tangent.

Tangency implies that

$$\frac{F}{n^2} - \left[u^{-1}\right]' (v(n) + \bar{u})\, v'(n) = 2/Z^2. \tag{A.8}$$

The left-hand side is strictly decreasing in n and satisfies $\lim_{n \to 0} = \infty$ and $\lim_{n \to \infty} = -\infty$, and thus has a unique solution that depends solely on $\bar{u}$ and parameters. □

Lemma 1. *Define* $\bar{n}$ *to be the population that maximizes the* $OL\,(n, x)$ *curve. Given outside option* $\bar{u}$*, the level of population* $\bar{n}$ *is unique and independent of location.*

Proof. From Proposition 1, the $OL\,(n, x)$ curve is strictly concave and twice continuously differentiable. The portion of the $OL\,(n, x)$ curve with $n(x) > 0$ is given:

$$OL\,(n, x) = y(x) + \Pi - \frac{F}{n} - u^{-1}\,(v(n) + \bar{u})\,.$$

The following first order condition is thus necessary and sufficient for the maximum:

$$\frac{F}{n^2} - \left[u^{-1}\right]'(v(n) + \bar{u})\, v'(n) = 0. \tag{A.9}$$

Call the population level that solves this $\bar{n}$. Given $\bar{u}$, the value of $\bar{n}$ is constant and independent of location. Write as $\mathcal{N}(\bar{u})$ the strictly decreasing function that gives the level of population that solves Equation A.9 as a function of $\bar{u}$. □

Lemma 2. *Define location $\hat{x}(\bar{u}, \Pi)$ as follows:*

$$\hat{x}(\bar{u}, \Pi) = \begin{cases} x \text{ s.t. } OL\left(\mathcal{N}(\bar{u}), \hat{x}\right) = MC\left(\mathcal{N}(\bar{u})\right) & \text{if } OL\left(\mathcal{N}(\bar{u}), 1\right) \geq MC\left(\mathcal{N}(\bar{u})\right) \\ 1 & \text{otherwise.} \end{cases}$$

Then $\hat{x}(\bar{u}, \Pi)$ is unique and in the range $[\underline{x}, 1]$.

Proof. In the first case, the highest-productivity location has an $OL(n, 1)$ curve with maximum value that exceeds the marginal cost at the corresponding level of population $\bar{n}$. At the threshold occupied location $\underline{x}$, the $OL(n, \underline{x})$ and $MC(n)$ curves meet with tangency. The marginal cost curve has positive slope, and so this tangency must occur to the left of the maximum of the $OL(n, \underline{x})$ curve. Evaluated at $\mathcal{N}(\bar{u})$, $OL(\mathcal{N}(\bar{u}), x)$ is continuous and strictly increasing in x. Therefore there exists a unique $\hat{x}(\bar{u}, \Pi) \in [\underline{x}, 1]$ such that $OL(\mathcal{N}(\bar{u}), \hat{x}(\bar{u}, \Pi)) = MC(\mathcal{N}(\bar{u}))$.

In the second case, $OL(\mathcal{N}(\bar{u}), x) < MC(\mathcal{N}(\bar{u}))$ for all locations. Then $\hat{x}(\bar{u}, \Pi)$ is uniquely defined as 1 and is within the specified range. □

Proposition 4. *In a stable general equilibrium, the following characterizes the local equilibrium populations and optimal zoning. For all locations $x > \hat{x}$, $n^*(x) = \bar{n}(x) = \mathcal{N}(\bar{u})$. For all locations $x \in [\underline{x}, \hat{x}]$, the population $n^*(x)$ and optimal zoning $\bar{n}(x)$ are given by the upper intersection of the $OL(n, x)$ and $MC(n)$ curves. For all $x \geq \underline{x}$, $p^*(x)$ is given by the value of the $OL(n, x)$ curve evaluated at $n^*(x)$.*

Proof. For locations $x > \hat{x}$, the $OL(n,x)$ curve is strictly greater than the $MC(n)$ curve, and therefore the equilibrium population can only be given by $\bar{n}$ if the initial residents choose the zoning law $\bar{n}(x) = \bar{n}$. I will show that this is the case. Note $OL(n,x) > MC(n)$ implies that the participation constraint will not bind, and so the solution to the initial resident problem is given by the first order condition

$$u'\left(y(x) + \Pi - \frac{F}{n} - p_0(x)\right)\frac{F}{n^2} = v'(n).$$

And in stable equilibrium, $p_0(x) = p(x)$ so substituting from the household spatial equilibrium condition gives

$$u'\left(u^{-1}\left(v(n) + \bar{u}\right)\right)\frac{F}{n^2} = v'(n).$$

From above, the maximum of the $OL(n,x)$ curve is given by the n that solves

$$\left[u^{-1}\right]'\left(v(n) + \bar{u}\right)v'(n) = \frac{F}{n^2}.$$

By the definition of the derivative of the inverse, these are equivalent: for locations with non-binding participation constraints, the level of population $\mathcal{N}(\bar{u})$ that gives the maximum of the $OL(n,x)$ curve is identical to that which maximizes initial resident utility (given the stability condition that $p_0(x) = p(x)$). For all locations $x > \hat{x}(\bar{u},\Pi)$, $\mathcal{N}(\bar{u})$ is consistent with the participation constraint and hence $n^*(x) = \bar{n}(x) = \mathcal{N}(\bar{u})$ for such locations. By Equation (1.6), the equilibrium price for such locations is given by $OL(\mathcal{N}(\bar{u}), x)$.

Consider location $x \in [\underline{x}, \hat{x}]$. From above, with $p_0(x) = p(x)$ the unconstrained choice of zoning that maximizes initial resident utility is the level of population that maximizes the $OL(n,x)$ curve. By definition, for locations with $x < \hat{x}(\bar{u},\Pi)$ the maximum of the $OL(n,x)$ curve is below the $MC(n)$ curve and hence infeasible. By concavity of the initial resident problem, the participation constraint will bind and the local equilibrium population (and optimal zoning law) is given by the upper intersection of the $OL(n,x)$

and $MC(n)$ curves. □

Proposition 5. *The utility-maximizing choices for the set of opened locations $\mathcal{X}$ and the set of zoning laws $\bar{n}(x)$ are identical to those chosen by the constrained planner. The equilibrium price gradient and aggregate profits implement the planner's choice of consumption. This set of instruments allows households to fully implement the planner's allocation.*

Proof. Under this voting regime, the measure 1 of households cast the decisive vote. As the vote takes place before development, households are identical and so the "median vote" is simply the representative household. For any set of zoning choices, the representative household rationally expects prices and consumption bundles to adjust to equalize utility across all occupied locations.

The representative household problem thus implies choosing the set of open locations and the set of zoning laws in order to maximize the level of utility $\bar{u}$ consistent with rational expectations about spatial equilibrium across all locations. Equivalently, the representative household can also choose consumption $c(x)$ for each location subject to the location-specific household budget constraint.

In making these choices, households are constrained by the spatial equilibrium conditions for all open locations, the budget constraints of households in all open locations, the definition of aggregate profits, the population constraint, and to a participation constraint.[1] The participation constraint implies that the price paid by households is consistent with firm construction decisions at each location.

As before, define $\theta(\bar{n}(x); \bar{u}, \Pi)$ as the equilibrium population for a location as a function of the zoning law for particular values of $\bar{u}$ and Π. With this notation, the represen-

[1]It could theoretically be possible that the utility-maximizing choice is to leave some households out of the urban economy, but by assumption the income of the agricultural sector is low enough that this is not the case.

tative household solves

$$\max_{\{\mathcal{X},\bar{n}(x),c(x),\bar{u}\}} \bar{u}$$

subject to the spatial equilibrium condition

$$u\left(c(x)\right) - v\left(\theta\left(\bar{n}(x);\bar{u},\Pi\right)\right) \leq \bar{u}$$

with equality if $\theta\left(\bar{n}(x);\bar{u},\Pi\right) > 0$ and for all $x \in \mathcal{X}$, subject to the household budget constraints

$$y(x) + \Pi - \frac{F}{\theta\left(\bar{n}(x);\bar{u},\Pi\right)} - p(x) - c(x) = 0,$$

subject to the definition of profits

$$\Pi = \int_{\mathcal{X}} p(x)\theta\left(\bar{n}(x);\bar{u},\Pi\right) - \left(\theta\left(\bar{n}(x);\bar{u},\Pi\right)/Z\right)^2 dx,$$

subject to the population constraint

$$1 = \int_{\mathcal{X}} \theta\left(\bar{n}(x);\bar{u},\Pi\right) dx,$$

and subject to the participation constraint

$$p(x) \geq 2\theta\left(\bar{n}(x);\bar{u},\Pi\right)/Z^2.$$

Now, note that adding a constant increment Δ to the price for every location raises aggregate profits by a total of Δ. Further, if the spatial equilibrium condition holds a particular set of prices and profits, it will also hold when after adding Δ to the price in each location. Spatial equilibrium only pins down the *relative* price gradient. If it is the case that zoning laws bind in all locations, this further implies that any price gradient and incumbent definition of aggregate profits that satisfies spatial equilibrium will not distort location choices. As before, if the participation constraint is not binding then

$\theta(\bar{n}(x); \bar{u}, \Pi) = \bar{n}(x)$. As any price gradient can be shifted upward by a sufficient increment such that the participation constraint does not bind and the remaining equations are unaffected, the representative agent's problem can thus be rewritten as a direct choice of the local population with no participation constraint.

Rewriting the budget constraint yields

$$p(x) = y(x) + \Pi - \frac{F}{n(x)} - c(x).$$

Substituting the full set of household budget constraints directly into the profit definition yields

$$\Pi = \int_{\mathcal{X}} \bar{n}(x) \left(y(x) + \Pi - \frac{F}{\bar{n}(x)} - c(x) \right) - \left(\frac{\bar{n}(x)}{Z} \right)^2 dx.$$

Rewriting, we have

$$0 = \int_{\mathcal{X}} \left[(y(x) - c(x))\, \bar{n}(x) - F - (\bar{n}(x)/Z)^2 \right] dx.$$

This is simply the aggregate resource constraint face by a constrained planner. Choosing a local zoning law $\bar{n}(x)$, consumption level $c(x)$, and outside option $\bar{u}$ for all locations x that satisfies the population and aggregate resource constraint implicitly defines a set of prices and level of profit consistent with the definition of aggregate profits.

Finally, the set $\mathcal{X}$ consists of the set $[\underline{x}, 1]$ for some $\underline{x}$. If not, there would exist an unoccupied location $\tilde{x}$ for which some $x < \tilde{x}$ were occupied. In this case, the representative household could loosen the aggregate resource constraint by changing the zoning in any occupied location $x < \tilde{x}$ from $\bar{n}(x)$ to 0 and in $\tilde{x}$ from 0 to $\bar{n}(x)$. Loosening the binding aggregate resource constraint must increase the achievable level of utility $\bar{u}$, and so the representative household will always make this choice.

With these considerations, the restated problem is now

$$\max_{\{n(x),c(x),\bar{u},\underline{x}\}} \bar{u}$$

subject to spatial equilibrium conditions

$$u\left(c(x)\right) - v\left(n(x)\right) = \bar{u},$$

the aggregate resource constraint

$$0 = \int_{\underline{x}}^{1} \left[\left(y(x) - c(x)\right) n(x) - F - \left(\frac{n(x)}{Z}\right)^{2} \right] dx,$$

and the population constraint

$$1 = \int_{\underline{x}}^{1} n(x) dx.$$

This problem is identical in objective and constraints to that of the constrained planner and so must share a solution. □

APPENDIX B

Appendix to *Entrepreneurship, Information, and Growth*

B.1 Data

Data on establishment dynamics comes from the Census Bureau's Business Information Tracking Series. Establishment births and deaths are divided by population data from the *Regional Economic Information System* to yield birth and death rates. In some regressions, we also include the de-meaned product of the birth and death rates. We utilize 1998–1999 birth and death data as the base period as it is the earliest with readily-available data.

The instrumental variables feature data drawn from the 1974 *County Business Patterns* from the Bureau of Census. The mining employment data was, in some cases, suppressed and only a range was reported. In these instances, total employment was estimated using the midpoint of the smallest range possible based on, first, the reported range for county mining employment and, second, the range implied by the counts of establishments in different employment size bins. A county with a reported mining employment range of 10-19 and two mining establishments each with employment in the range of 6-10 would have a narrowed range of 12-19, and an estimated employment of 16.5. When calculating the log of mining employment, the final employment counts and estimates were increased by one employee to accommodate counties with no employment.

The other measures of entrepreneurship utilized in the robustness checks are based on proprietorship and the share of employment in businesses with fewer than five employees. The two proprietorship measures are the share of proprietors in total employment in 1998 and the growth rate of proprietors from 1988 to 1998 weighted by 1988 employment. Data on proprietors is from the *Regional Economic Information System* of the Bureau of Census, as is the total employment data. The share of employment in small businesses with 1 to 4 employees during 1998 was estimated using *County Business Patterns* data from the Bureau of Census. In IV results using these measures, lagged values of these variables from, respectively, 1969, the period 1969–1979, and 1974 were used as instruments alongside the mining and density variables.

The key dependent variable is the rate of growth in non-farm employment between 2000 and 2007, a peak-to-peak period in terms of the business cycle. Robustness results include growth from 2000–2011. Data on the growth variables for the relevant years come from the *Regional Economic Information System* produced by the Bureau of Economic Analysis of the U.S. Department of Commerce.

The data for occupations in the arts and requiring a high degree of human capital are taken from the USDA's Economic Research Service as explained in the text. The population and the share of the population with a bachelor's degree (including those with additional degrees) is taken from the 2000 Census.

Tract-weighted population density for the year 2000 is constructed using tract-level data on population and land area. Tract population density is weighted by tract population and summed to the county level. The tract-weighted measure is included as traditional population density measures may be heavily influenced by the non-urban area of the county, which is quite heterogeneous: Los Angeles County is much larger than the city and contains portions of multiple national forests, whereas Manhattan is its own county. The key results are similar when using either measure.

Log income in 2000, log employment in 2000, and employment growth from 1990-2000 are constructed using REIS data. The local demand shocks for 2000–2007 (for most regressions) and 2000–2011 (for robustness regressions) use a combination of REIS data for government employment with CBP data for most industries. The local demand shock variables are constructed using employment totals for primarily three-digit NAICS codes, as described in the text. For suppressed values, we implement a simplified version of the technique outlined in Isserman and Westervelt (2006) that is similar to that described above for the mining instruments.

Median age in 2000 is taken from Census Bureau estimates. Population growth in the period 1950–1960 is taken from the decennial censuses. The amenity scores are the standardized scores generated by Economic Research Service of the United States Department of Agriculture. For the distance measures, MSA populations and distances are taken from and constructed using county populations in the 2000 Census.

The analysis in the paper focuses on two geographic levels: counties (and county-equivalents) and metropolitan statistical areas. The possibility of geographic information asymmetries encourages a focus on local areas that are likely to contain the informational spillovers generated by entrepreneurial or research activity. Counties are the smallest geographic unit for which the relevant data—notably on establishment births and deaths, but other control variables as well—is readily available, and so we focus our analysis there. Metropolitan areas are also included as inter-county commuting ties suggest that information may flow across these jurisdictional boundaries.

The United States is divided into 3,143 counties and county equivalents, including a number of independent cities that are economically integrated with their surrounding counties. Upon aggregation of independent cities with surrounding counties, 3,111 counties remain. Of these, we focus on the lower forty-eight states due to concerns about data and applicability. In addition, there are a number of instances of county formations during the period in question; these counties are aggregated to their largest extent in order to ensure data consistency. This leaves a sample of 3,072 counties, county equivalents, and county aggregates that encompass the entirety of the continental United States. When aggregating these finalized counties to the relevant metropolitan areas according to year 2000 definitions, we end up with 356 metropolitan areas.

B.2 Tables

Birth Rate, 1998	0.87***	0.7***	0.68***	0.62***
	(0.025)	(0.038)	(0.036)	(0.09)
Death Rate, 1998	—	0.23***	0.23***	0.35***
		(0.039)	(0.04)	(0.1)
B*D, 1998	—	0.036**	0.01	0.038
		(0.015)	(0.014)	(0.026)
Demand Shock, 2000–07	0.007	-0.001	0.004	-0.001
	(0.007)	(0.006)	(0.006)	(0.019)
Emp Growth, 1990-2000	0.007***	0.009***	0.01***	0.012***
	(0.002)	(0.002)	(0.002)	(0.003)
Log Employment, 2000	0.11***	0.11***	0.09***	0.05
	(0.029)	(0.029)	(0.025)	(0.05)
Log Income, 2000	-0.79***	-0.79***	-0.57***	-0.3
	(0.162)	(0.159)	(0.14)	(0.3)
Density, 2000	0.004	0.003	-0.004	-0.001
	(0.004)	(0.003)	(0.003)	(0.005)
HC Share, 2000	0.030**	0.032***	0.04***	0.022
	(0.013)	(0.012)	(0.012)	(0.028)
Arts Share, 2000	-0.028	-0.026	0.035	-0.07
	(0.078)	(0.076)	(0.07)	(0.12)
BA Share, 2000	-0.012	-0.011	-0.024***	-0.016
	(0.008)	(0.007)	(0.007)	(0.018)
Pop Growth, 1950-1960	0.047	0.017	-0.1	-0.23***
	(0.060)	(0.07)	(0.06)	(0.07)
Median Age, 2000	0.016**	0.007	-0.005	-0.005
	(0.007)	(0.007)	(0.008)	(0.019)
Amenity Score	0.043***	0.042***	0.024*	-0.034
	(0.008)	(0.008)	(0.012)	(0.031)
Distance to MSA	0.002***	0.003***	0.002***	0.009
	(0.001)	(0.001)	(0.001)	(0.007)
Marginal distance to MSA > 250k	-0.001	0	0	-0.009
	(0.0001)	(0.001)	(0.000)	(0.008)
Marginal distance to MSA > 500k	0.002***	0.002***	0	-0.001
	(0.001)	(0.001)	(0.000)	(0.006)
Marginal distance to MSA > 1M	-0.001***	-0.001***	-0.001**	0.005
	(0.000)	(0.000)	(0.000)	(0.004)
Constant	6.66***	6.69***	4.89***	3.81
	(1.55)	(1.54)	(1.37)	(3.44)
N	3072	3072	3072	825
R^2	0.78	0.79	0.82	0.95
F (non-FEs)	272	242	89.9	27.5
Counties	All	All	All	Metro
Fixed Effects	No	No	State	MSA
Full Controls	Yes	Yes	Yes	Yes
Regression	OLS	OLS	OLS	OLS

Table B.1: Preliminary Entrepreneurship Results. Dependent variable is the establishment birth rate, 2005. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 3.1 and the data appendix for details on the variables.

Log Mining Emp, 1974	-0.08***	0.016	0.034	-0.1
	(0.022)	(0.037)	(0.034)	(0.06)
Mining Emp > 20	—	-0.12	-0.14	-0.1
		(0.12)	(0.1)	(0.2)
Mining Emp > 100	—	-0.24**	-0.12	-0.2
		(0.11)	(0.1)	(0.22)
Density, 1980	—	0.1	0.09	0.31**
		(0.07)	(0.07)	(0.14)
Lagged Mining×Density	—	-0.011***	-0.011***	-0.02***
		(0.003)	(0.003)	(0.006)
Predicted Employment Growth, 2000–07	0.08***	0.07***	0.07***	0.05***
	(0.009)	(0.009)	(0.009)	(0.018)
Lagged Employment Growth, 1990-2000	0.014***	0.014***	0.003	0.008*
	(0.003)	(0.003)	(0.002)	(0.005)
Log Employment, 2000	0.11*	0.035	-0.043	0.022
	(0.05)	(0.05)	(0.042)	(0.09)
Log Income, 2000	-1.8***	-1.61***	-1.11***	-0.28
	(0.33)	(0.28)	(0.27)	(0.54)
Density, 2000	0.01	-0.005	0.000	-0.09
	(0.007)	(0.029)	(0.03)	(0.06)
HC Share, 2000	0.08***	0.09***	0.08***	-0.06
	(0.022)	(0.022)	(0.02)	(0.046)
Arts Share, 2000	0.047	0.28**	0.27**	0.55**
	(0.12)	(0.12)	(0.11)	(0.25)
BA Share, 2000	-0.03***	-0.037***	-0.045***	-0.015
	(0.014)	(0.014)	(0.012)	(0.025)
Pop Growth, 1950-1960	0.26**	0.3**	0.29*	0.6
	(0.13)	(0.14)	(0.15)	(0.54)
Median Age, 2000	0.08***	0.08***	0.08***	0.11***
	(0.011)	(0.011)	(0.009)	(0.023)
Amenity Score	0.1***	0.08***	0.07***	0.012
	(0.012)	(0.014)	(0.014)	(0.031)
Distance to MSA	0.004***	0.004***	0.002***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Marginal distance to MSA > 250k	0.000	0.000	0.000	-0.002**
	(0.000)	(0.000)	(0.000)	(0.001)
Marginal distance to MSA > 500k	0.001*	0.001*	0.001	0.002**
	(0.001)	(0.001)	(0.000)	(0.001)
Marginal distance to MSA > 1M	-0.001***	-0.001**	0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)
Constant	17.2***	15.4***	11.3***	0.19
	(3.11)	(2.7)	(2.53)	(5.15)
N	3072	3072	3072	3072
R^2	0.49	0.52	0.44	0.16
F—Regression	42.1	37.6	37.9	5.7
F—IV	13.9	8.87	8.26	6.86
Dependent Variable	Births	Births	Deaths	Births×Deaths
Counties	All	All	All	All
Fixed Effects	No	No	No	No
Regression	OLS	OLS	OLS	OLS

Table B.2: First Stage Results. Dependent variable is the establishment birth rate, 1998. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 3.1 and the data appendix for details on the variables.

Birth Rate, 1998	—	3.79***	—	3.32*
		(0.36)		(1.72)
Proprietorship Share, 1998	-0.12	-0.13	-0.14	-0.6**
	(0.14)	(0.13)	(0.32)	(0.3)
Proprietorship Growth Rate, 1988–98	0.63***	0.67***	4.02	1.4
	(0.17)	(0.14)	(2.56)	(1.49)
Small Business Employment Share, 1998	0.7***	-0.31	-1.46	-0.13
	(0.19)	(0.19)	(1.18)	(0.81)
Predicted Employment Growth, 2000–07	0.81***	0.69***	1.29***	0.76**
	(0.12)	(0.11)	(0.36)	(0.32)
Lagged Employment Growth, 1990–2000	0.18***	0.12***	-0.27	0.043
	(0.047)	(0.038)	(0.35)	(0.18)
Log Employment, 2000	0.13	-0.33	-0.17	-0.26
	(0.45)	(0.42)	(0.76)	(0.44)
Log Income, 2000	-13.3***	-10.1***	0.98	-10.8
	(3.94)	(3.5)	(12.7)	(8.14)
Density, 2000	0.09**	0.08**	0.04	0.06
	(0.037)	(0.035)	(0.08)	(0.045)
HC Share, 2000	0.91***	0.67***	0.3	0.74**
	(0.27)	(0.25)	(0.55)	(0.36)
Arts Share, 2000	-1.78*	-1.96**	-2.62**	-2.07**
	(0.93)	(0.86)	(1.33)	(0.91)
BA Share, 2000	-0.42***	-0.33**	-0.14	-0.34*
	(0.14)	(0.13)	(0.27)	(0.18)
Population Growth, 1950-1960	1.99**	0.93	4.18*	1.28
	(0.91)	(0.78)	(2.23)	(1.72)
Median Age, 2000	-0.35***	-0.52***	-0.23	-0.44***
	(0.11)	(0.1)	(0.16)	(0.14)
Amenity Score	0.2	0.016	-0.49	0.09
	(0.15)	(0.14)	(0.66)	(0.37)
Distance to MSA	-0.017*	-0.018**	-0.019	-0.017*
	(0.009)	(0.008)	(0.015)	(0.009)
Marginal distance to MSA > 250k	-0.004	-0.001	0.006	0.003
	(0.004)	(0.004)	(0.009)	(0.005)
Marginal distance to MSA > 500k	-0.002	-0.004	-0.003	-0.004
	(0.004)	(0.004)	(0.006)	(0.004)
Marginal distance to MSA > 1M	0.001	0.004	0.009	0.003
	(0.004)	(0.003)	(0.008)	(0.005)
Constant	136***	109***	-4.54	118
	(38)	(33.4)	(127)	(81.7)
N	3072	3072	3072	3072
R^2	0.53	0.58	—	0.56
F	49.1	55.6	17.6	45.4
Counties	All	All	All	All
Fixed Effects	No	No	No	No
Regression	OLS	OLS	IV	IV
Weak ID F-test			4.72	1.25
Weak IV-robust p-value	—	—	0.5	0.046
Endogeneity of Entrepreneurship p-value	—	—	0.001	0.005
Endogeneity of IVs p-value			—	0.93

Table B.3: Robustness. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 3.1 and the data appendix for details on the variables. The weak instrument-robust p-value is for the Anderson-Wald F test statistic; the weak instrument test is the Cragg-Donald Wald F statistic. The test for the endogeneity of entrepreneurship (i.e., the birth rate) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J statistic.

Birth Rate, 1998	4.73***	5.2***	5.37***	7.74*	7.31*
	(0.56)	(0.75)	(0.75)	(4.32)	(4.35)
Death Rate, 1998	-1.84***	-3.2***	-2.97***	-9.44*	-6.11
	(0.5)	(0.84)	(0.84)	(5.07)	(4.64)
Births×Deaths, 1998	0.17	0.71**	0.49**	1.99**	1.05
	(0.19)	(0.28)	(0.22)	(0.94)	(0.84)
Predicted Employment Growth, 2000–07	0.67***	0.83***	0.68***	0.93***	0.64***
	(0.11)	(0.18)	(0.16)	(0.18)	(0.12)
Lagged Employment Growth, 1990–2000	0.19***	0.24***	0.26***	0.15**	0.16**
	(0.038)	(0.032)	(0.032)	(0.07)	(0.07)
Log Employment, 2000	-0.31	-0.11	-0.08	-0.35	-0.15
	(0.42)	(0.49)	(0.51)	(0.49)	(0.45)
Log Income, 2000	-12.5***	-14.7***	-16.2***	-16.1***	-15.2***
	(3.39)	(4.01)	(3.51)	(4.66)	(4.13)
Density, 2000	0.1***	0.11**	0.06	0.13***	0.06
	(0.035)	(0.042)	(0.04)	(0.043)	(0.039)
HC Share, 2000	0.81***	0.63**	0.45*	1.26***	0.92***
	(0.28)	(0.31)	(0.25)	(0.31)	(0.3)
Arts Share, 2000	-1.83**	-1.22	0.54	-2.01**	-0.43
	(0.89)	(1.06)	(0.96)	(1.02)	(0.9)
BA Share, 2000	-0.41***	-0.35*	-0.3**	-0.67***	-0.52***
	(0.14)	(0.18)	(0.14)	(0.17)	(0.14)
Pop Growth, 1950-1960	0.67	-0.48	-0.013	0.7	0.25
	(0.84)	(0.83)	(0.77)	(0.99)	(0.85)
Median Age, 2000	-0.5***	-0.36***	-0.53***	-0.33**	-0.62***
	(0.1)	(0.13)	(0.13)	(0.16)	(0.12)
Amenity Score	0.15	0.17	-0.22	0.47**	-0.16
	(0.14)	(0.15)	(0.27)	(0.18)	(0.25)
Distance to MSA	-0.017*	0.06***	-0.048**	0.000	-0.005
	(0.009)	(0.023)	(0.022)	(0.005)	(0.01)
Marginal distance to MSA > 250k	-0.003	-0.008	-0.02**	-0.004	-0.011*
	(0.004)	(0.005)	(0.008)	(0.005)	(0.006)
Marginal distance to MSA > 500k	-0.004	0.000	0.005	0.005	-0.001
	(0.004)	(0.005)	(0.007)	(0.005)	(0.005)
Marginal distance to MSA > 1M	0.004	0.006	0.003	0.408	0.006
	(0.003)	(0.004)	(0.005)	(0.422)	(0.005)
Constant	132***	150***	176***	175***	-173***
	(32.5)	(37.8)	(33.7)	(45.3)	(40.2)
N	3072	825	825	3072	3072
R^2	0.56	0.65	0.74	0.44	0.6
F	59.8	50.6	30.7	41.5	35.6
Counties	All	Metro	Metro	All	All
Fixed Effects	No	No	No	No	Yes
Regression	OLS	OLS	OLS	IV	IV
Weak ID F-test				5.98	7.86
Weak IV-Robust P-value	—	—	—	0.046	0.13
Endogeneity of Entrepreneurship P-value	—	—	—	0.5	0.99
Endogeneity of IVs P-value				0.99	0.9

Table B.4: Full Entrepreneurial Information Results. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 3.1 and the data appendix for details on the variables. The weak instrument-robust p-value is for the Anderson-Wald F test statistic; the weak instrument test is the Cragg-Donald Wald F statistic. The test for the endogeneity of entrepreneurship (i.e., the birth rate) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J statistic.

References

Ács, Z. and Armington, C. (2006). *Entrepreneurship, Geography, and American Economic Growth*. Cambridge University Press.

Ács, Z., Braunerhjelm, P., Audretsch, D., and Carlsson, B. (2009). The knowledge spillover theory of entrepreneurship. *Small Business Economics*, 32(1):15–30.

Akerlof, G. (1970). The market for 'lemons': Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, pages 488–500.

Albouy, D. (2012). Evaluating the efficiency and equity of federal fiscal equalization. *Journal of Public Economics*, 96(9):824–839.

Albouy, D. and Ehrlich, G. (2012). Metropolitan land values and housing productivity. Unpublished.

Albouy, D., Graf, W., Kellogg, R., and Wolff, H. (2013). Climate Amenities, Climate Change, and American Quality of Life. Working Paper 18925, National Bureau of Economic Research.

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 1085:1139.

Allen, T., Arkolakis, C., and Li, X. (2015). Optimal City Structure. Unpublished.

Arnott, R. J. and Stiglitz, J. E. (1979). Aggregate land rents, expenditure on public goods, and optimal city size. *The Quarterly Journal of Economics*, pages 471–500.

Audretsch, D. and Keilbach, M. (2004). Entrepreneurship, Capital and Economic Performance. *Regional Studies*, 38(8):949–959.

Audretsch, D. and Keilbach, M. (2007). The localisation of entrepreneurship capital: Evidence from germany. *Papers in Regional Science*, 86(3):351–365.

Aura, S. and Davidoff, T. (2008). Supply Constraints and Housing Prices. *Economics Letters*, 99(2):275–277.

Avent, R. (2011). *The Gated City.* Amazon Digital Services.

Bajari, P. and Benkard, C. L. (2005). Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy*, 113(6):1239–1276.

Barro, R. J. (2006). Rare Disasters and Asset Markets in the Twentieth Century. *The Quarterly Journal of Economics*, 121(3):823–866.

Barro, R. J. (2013). Environmental Protection, Rare Disasters, and Discount Rates. Working Paper 19258, National Bureau of Economic Research.

Basu, N. and Chevrier, M. (2011). Distance, information asymmetry, and mergers: Evidence from canadian firms. *Managerial Finance*, 37(1):21–33.

Beeson, P., DeJong, D., and Troesken, W. (2001). Population growth in us counties, 1840–1990. *Regional Science and Urban Economics*, 31(6):669–699.

Behrens, K., Mion, G., Murata, Y., and Südekum, J. (2014). Spatial frictions. Technical report, Düsseldorf Institute for Competition Economics.

Blanchard, O., Katz, L., Hall, R., and Eichengreen, B. (1992). Regional evolutions. *Brookings papers on economic activity*, 1992(1):1–75.

Braunerhjelm, P., Ács, Z. J., Audretsch, D. B., and Carlsson, B. (2010). The missing link: Knowledge diffusiong and entrepreneurship in endogenous growth. *Small Business Economics*, 34(2):105–125.

Bunten, D., Weiler, S., Thompson, E., and Zahran, S. (2015). Entrepreneurship, Information, and Growth. *Journal of Regional Science*, 55(4):560–584.

Calabrese, S., Epple, D., and Romano, R. (2007). On the Political Economy of Zoning. *Journal of Public Economics*, 91(1):25–49.

Chay, K. Y. and Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, 113(2):376–424.

Chen, C.-S. and Lai, F.-C. (2008). Location choice and optimal zoning under cournot competition. *Regional Science and Urban Economics*, 38(2):119–126.

Cheshire, P. C. and Hilber, C. A. (2008). Office Space Supply Restrictions in Britain: The Political Economy of Market Revenge. *The Economic Journal*, 118(529):F185–F221.

Chinitz, B. (1961). Contrasts in Agglomeration: New York and Pittsburgh. *American Economic Review*, 51(2):279–289.

Costello, C. J., Neubert, M. G., Polasky, S. A., and Solow, A. R. (2010). Bounded Uncertainty and Climate Change Economics. *Proceedings of the National Academy of Sciences*, 107(18):8108–8110.

Costinot, A., Donaldson, D., and Smith, C. (2012). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from a 9 million-field partition of the earth. *Manuscript*.

Davis, D. R. and Dingel, J. I. (2014). The Comparative Advantage of Cities. Technical report, National Bureau of Economic Research.

Deller, S., Tsai, T., Marcouiller, D., and English, D. (2001). The role of amenities and quality of life in rural economic growth. *American Journal of Agricultural Economics*, 83(2):352–365.

Desmet, K. and Rossi-Hansberg, E. (2013a). On the Spatial Economic Impact of Global Warming. Manuscript.

Desmet, K. and Rossi-Hansberg, E. (2013b). Urban accounting and welfare. *American Economic Review*, 103(6):2296–2327.

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000. *American Economic Review*, 106(3):479–524.

Eeckhout, J. and Guner, N. (2015). Optimal Spatial Taxation: Are Big Cities Too Small? Technical report, Institute for the Study of Labor.

Ehrlich, I. and Becker, G. S. (1972). Market Insurance, Self-Insurance, and Self-Protection. *The Journal of Political Economy*, pages 623–648.

Fischel, W. A. (1987). *The Economics of Zoning Laws: A Property Rights Approach to American Land Use Controls*. Johns Hopkins University Press.

Fischel, W. A. (1999). Does the American Way of Zoning Cause the Suburbs of Metropolitan Areas to be Too Spread Out? *Governance and Opportunity in Metropolitan America*, 15:1–19.

Fischel, W. A. (2009). *The Homevoter Hypothesis*. Harvard University Press.

Fischel, W. A. (2015). *Zoning Rules!* Lincoln Institute of Land Policy, Cambridge, Massachusetts.

Flatters, F., Henderson, V., and Mieszkowski, P. (1974). Public goods, efficiency, and regional fiscal equalization. *Journal of Public Economics*, 3(2):99–112.

Gamper-Rabindran, S. and Timmins, C. (2011). Hazardous Waste Cleanup, Neighborhood Gentrification, and Environmental Justice: Evidence from Restricted Access Census Block Data. *The American Economic Review*, 101(3):620–624.

Ganong, P. and Shoag, D. (2013). Why has Regional Income Convergence in the US Declined. *Manuscript, Harvard.*

Glaeser, E., Kerr, S., and Kerr, W. (2012). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Working*.

Glaeser, E., Scheinkman, J., and Shleifer, A. (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics*, 36(1):117–143.

Glaeser, E. L. and Gyourko, J. (2002). Zoning's Steep Price. *Regulation*, 25:24.

Glaeser, E. L. and Gyourko, J. (2003). The Impact of Building Restrictions on Housing Affordability. *Federal Reserve Bank of New York. Economic Policy Review*, 9(2):21.

Glaeser, E. L. and Gyourko, J. (2005). Urban Decline and Durable Housing. *Journal of Political Economy*, 113(2):345–375.

Glaeser, E. L., Gyourko, J., and Saks, R. E. (2005a). Why Have Housing Prices Gone Up? *The American Economic Review*, 95(2):pp. 329–333.

Glaeser, E. L., Gyourko, J., and Saks, R. E. (2005b). Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices. *Journal of Law and Economics*, 48(2):331–369.

Glaeser, E. L., Gyourko, J., and Saks, R. E. (2006). Urban Growth and Housing Supply. *Journal of Economic Geography*, 6(1):71–89.

Goodell, J. (2013). *Goodbye, Miami.* http://www.rollingstone.com/politics/news/why-the-city-of-miami-is-doomed-to-drown-20130620. [Online; accessed 3/21/2014].

Graves, P., Weiler, S., and Tynon, E. (2009). The economics of ghost towns. *Journal of Regional Analysis and Policy*, 39(2):131–140.

Greenstone, M. and Gallagher, J. (2008). Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program. *The Quarterly Journal of Economics*, 123(3):951–1003.

Gyourko, J., Mayer, C., and Sinai, T. (2013). Superstar Cities. *American Economic Journal: Economic Policy*, 5(4):167–199.

Gyourko, J. and Molloy, R. (2014). Regulation and Housing Supply. Technical report, National Bureau of Economic Research.

Hall, R., Gordon, R., and Holt, C. (1970). Why is the unemployment rate so high at full employment? *Brookings Papers on Economic Activity*, 1970(3):369–410.

Hall, R. E. and Jones, C. I. (2007). The Value of Life and the Rise in Health Spending. *The Quarterly Journal of Economics*, 122(1):39–72.

Hamilton, B. W. (1975). Zoning and Property Taxation in a System of Local Governments. *Urban Studies*, 12(2):205–211.

Hammond, G. and Thompson, E. (2008). Determinants of income growth in metropolitan and nonmetropolitan labor markets. *American Journal of Agricultural Economics*, 90(3):783–793.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding Instrumental Variables in Models with Essential Heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432.

Henderson, J. V. (1991). Optimal regulation of land development through price and fiscal controls. *Journal of Urban Economics*, 30(1):64–82.

Hilber, C. A. and Robert-Nicoud, F. (2013). On the origins of land use regulations: Theory and evidence from us metro areas. *Journal of Urban Economics*, 75:29–43.

Hills, R. and Schleicher, D. (2011). Balancing the "Zoning Budget". *Case Western Reserve Law Review*, 42(81):112–118.

Hills, R. M. and Schleicher, D. (2014). City Replanning. *George Mason Law & Economics Research Paper*, (14-32).

Hochman, O., Pines, D., and Thisse, J.-F. (1995). On the optimal structure of local governments. *The American Economic Review*, pages 1224–1240.

Hsieh, C.-T. and Moretti, E. (2015). Why Do Cities Matter? Local Growth and Aggregate Growth. Technical report, National Bureau of Economic Research.

Huang, T., Orazem, P., and Wohlgemuth, D. (2002). Rural population growth, 1950–1990: the roles of human capital, industry structure, and government policy. *American Journal of Agricultural Economics*, 84(3):615–627.

Isserman, A. M. and Westervelt, J. (2006). 1.5 million missing numbers: Overcoming employment suppression in county business patterns data. *International Regional Science Review*, 29(3):311–335.

Jacobs, J. (1970). *The Economy of Cities*. Vintage, New York.

Kahn, M. E. (2005). The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions. *The Review of Economics and Statistics*, 87(2):pp. 271–284.

Kahn, M. E. (2009). Urban Growth and Climate Change. *Annual Review of Resource Econonomics*, 1(1):333–350.

Kanemoto, Y. (1988). Hedonic Prices and the Benefits of Public Projects. *Econometrica*, pages 981–989.

Knabb, R. D., Rhome, J. R., and Brown, D. P. (2005). *Tropical Cyclone Report. Hurricane Katrina*. http://www.nhc.noaa.gov/pdf/TCR-AL122005_Katrina.pdf. [Online; accessed 3/21/2014].

Koster, H. R., van Ommeren, J., and Rietveld, P. (2012). Bombs, Boundaries and Buildings: a Regression-Discontinuity Approach to Measure Costs of Housing Supply Restrictions. *Regional Science and Urban Economics*, 42(4):631–641.

Kousky, C., Luttmer, E. F., and Zeckhauser, R. J. (2006). Private Investment and Government Protection. *Journal of Risk and Uncertainty*, 33(1-2):73–100.

Krugman, P. (1987). The narrow moving band, the dutch disease, and the competitive consequences of mrs. thatcher: Notes on trade in the presence of dynamic scale economies. *Journal of development Economics*, 27(1):41–55.

Krugman, P. (1991). Increasing Returns and Economic Geography. *The Journal of Political Economy*, 99(3):483–499.

Krupka, D. J. (2009). Location-Specific Human Capital, Location Choice, and Amenity Demand. *Journal of Regional Science*, 49(5):833–854.

Kuminoff, N. V. and Pope, J. C. (Forthcoming). Do 'capitalization effects' for public goods reveal the public's willingness to pay? *International Economic Review.*

Kuminoff, N. V., Smith, V. K., and Timmins, C. (2013). The new economics of equilibrium sorting and policy evaluation using housingmarkets. *Journal of Economic Literature*, 51(4):1007–1062.

Lang, W. and Nakamura, L. (1993). A model of redlining. *Journal of Urban Economics*, 33(2):223–234.

Lind, R. C. (1973). Spatial Equilibrium, the Theory of Rents, and the Measurement of Benefits from Public Programs. *The Quarterly Journal of Economics*, 87(2):188–207.

Low, S. A. and Weiler, S. (2012). Employment risk, returns, and entrepreneurship. *Economic Development Quarterly*, 26(3):238–251.

Lucas, R. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1):3–42.

Mayer, C. J. and Somerville, C. T. (2000). Land Use Regulation and New Construction. *Regional Science and Urban Economics*, 30(6):639–662.

Miami-Dade County (2010). *GreenPrint*. http://www.miamidade.gov/greenprint/pdf/plan.pdf. [Online; accessed 3/21/2014].

Monkkonen, P. and Quigley, J. M. (2008). The Spatial Consequences of Autarky in Land-Use Regulation: Strategic Interaction or Parallelism? *Berkeley Program on Housing and Urban Policy*.

Morales, E., Serrato, J. C., and Zidar, O. (2015). State Taxes and Spatial Misallocation. Working Paper.

Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121:175–212.

Moretti, E. (2011). Local labor markets. *Handbook of Labor Economics*, 4:1237–1313.

Mulvihill, G. (2013). *Tallying Hurricane Sandy Deaths and Damage a Complex Task for Officials*. http://www.huffingtonpost.com/2013/10/25/hurricane-sandy-deaths_n_4164333.html. [Online; accessed 3/21/2014].

Newman, A. (2012). *Hurricane Sandy vs. Hurricane Katrina*. http://cityroom.blogs.nytimes.com/2012/11/27/hurricane-sandy-vs-hurricane-katrina/?_php=true&_type=blogs&_r=0. [Online; accessed 3/21/2014].

Nordhaus, W. D. and Yang, Z. (1996). A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies. *American Economic Review*, 86(4):741–765.

Ortalo-Magné, F. and Prat, A. (2014). On the Political Economy of Urban Growth: Homeownership versus Affordability. *American Economic Journal: Microeconomics*, 6(1):154–81.

Partridge, M., Rickman, D., Ali, K., and Olfert, R. (2009). Do new economic geography agglomeration shadows underlie current population dynamics across the urban hierarchy. *Papers in Regional Science*, 88(2):445–466.

Pindyck, R. S. (2013). Climate Change Policy: What Do the Models Tell Us? *Journal of Economic Literature*, 51(3):860–872.

Pindyck, R. S. and Wang, N. (2013). The Economic and Policy Consequences of Catastrophes. *American Economic Journal: Economic Policy*, 5(4):306–39.

Rappaport, J. (2004). Why are population flows so persistent? *Journal of Urban Economics*, 56(3):554–580.

Rappaport, J. (2007). Moving to nice weather. *Regional Science and Urban Economics*, 37(3):375–398.

Rappaport, J. and Sachs, J. D. (2003). The United States as a Coastal Nation. *Journal of Economic Growth*, 8(1):5–46.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *The Journal of Political Economy*, 90(6):1257–12T8.

Romer, P. (1986). Increasing returns and long-run growth. *The Journal of Political Economy*, pages 1002–1037.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5).

Rosen, S. (1979). Wage-Based Indexes of Urban Quality of Life. *Current Issues in Urban Economics*, 3.

Rosen, S. (2002). Markets and Diversity. *American Economic Review*, 92(1):1–15.

Rossi-Hansberg, E. (2004). Optimal Urban Land Use and Zoning. *Review of Economic Dynamics*, 7(1):69–106.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Schleicher, D. (2013). City unplanning. *Yale Law Journal*, 122(7):1670–1737.

Shapiro, J. (2006). Smart cities: Quality of life, productivity, and the growth effects of human capital. *The Review of Economics and Statistics*, 88(2):324–335.

Shogren, J. and Stamland, T. (2002). Skill and the Value of Life. *Journal of Political Economy*, 110(5):1168–1173.

Shogren, J. F. and Crocker, T. D. (1991). Risk, Self-Protection, and Ex Ante Economic Value. *Journal of Environmental Economics and Management*, 20(1):1–15.

Shogren, J. F. and Stamland, T. (2005). Self-Protection and Value of Statistical Life Estimation. *Land Economics*, 81(1):100–113.

Siodla, J. (2015). Razing San Francisco: The 1906 disaster as a natural experiment in urban redevelopment. *Journal of Urban Economics*, 89:48 – 61.

Starrett, D. A. (1981). Land Value Capitalization in Local Public Finance. *The Journal of Political Economy*, pages 306–327.

Stephens, H. and Partridge, M. (2011). Do entrepreneurs enhance economic growth in lagging regions? *Growth and Change*, 42(4):431–465.

Stephens, H. M., Partridge, M. D., and Faggian, A. (2013). Innovation, entrepreneurship and economic growth in lagging regions. *Journal of Regional Science*.

Stull, W. J. (1974). Land Use and Zoning in an Urban Economy. *American Economic Review*, pages 337–347.

Summers, L. H., Abraham, K. G., and Wachter, M. L. (1986). Why is the unemployment rate so very high near full employment? *Brookings Papers on Economic Activity*, 1986(2):339–396.

Turner, M. A., Haughwout, A., and van der Klaauw, W. (2014). Land Use Regulation and Welfare. *Econometrica*, 82(4):1341–1403.

Uysal, V., Kedia, S., and Panchapagesan, V. (2008). Geography and acquirer returns. *Journal of Financial Intermediation*, 17(2):256–275.

Van Landingham, M. J. (2007). Murder Rates in New Orleans, La, 2004–2006. *American Journal of Public Health*, 97(9):1614–1616.

Van Nieuwerburgh, S. and Weill, P.-O. (2010). Why Has House Price Dispersion Gone Up? *The Review of Economic Studies*, 77(4):1567–1606.

Weiler, S. (1997). The economics of the struggling structurally unemployed. *Journal of Appalachian Studies*, 3(1):71–98.

Weiler, S. (2000a). Pioneers and settlers in lo-do denver: Private risk and public benefits in urban redevelopment. *Urban Studies*, 37(1):167–179.

Weiler, S. (2000b). Unemployment in regional labor markets: Using structural theories to understand local jobless rates in west virginia. *Industrial and Labor Relations Review*, 54:573.

Weiler, S. (2001). Unemployment in regional labor markets: Using structural theories to understand local jobless rates in west virginia. *Industrial and Labor Relations Review*, pages 573–592.

Weiler, S. (2006). A park by any other name: National park designation as a natural experiment in signaling. *Journal of Urban Economics*, 60(1):96–106.

Weiler, S., Hoag, D., and Fan, C. (2006). Prospecting for economic returns to research: Adding informational value at the market fringe. *Journal of Regional Science*, 46(2):289–311.

Weitzman, M. L. (2009). On Modeling and Interpreting the Economics of Catastrophic Climate Change. *The Review of Economics and Statistics*, 91(1):1–19.

Weitzman, M. L. (2011). Fat-Tailed Uncertainty in the Economics of Catastrophic Climate Change. *Review of Environmental Economics and Policy*, 5(2):275–292.

Zheng, S. and Kahn, M. E. (2013). Does Government Investment in Local Public Goods Spur Gentrification? Evidence from Beijing. *Real Estate Economics*, 41(1):1–28.