American Economic Review 2020, 110(3): 720–759
https://doi.org/10.1257/aer.20141772

Segmented Housing Search[†]

By Monika Piazzesi, Martin Schneider, and Johannes Stroebel*

We study housing markets with multiple segments searched by heterogeneous clienteles. In the San Francisco Bay Area, search activity and inventory covary negatively across cities, but positively across market segments within cities. A quantitative search model shows how the endogenous flow of broad searchers to high-inventory segments within their search ranges induces a positive relationship between inventory and search activity across segments with a large common clientele. The prevalence of broad searchers shapes the response of housing markets to localized supply and demand shocks. Broad searchers help spread shocks across many segments and reduce their effect on local market activity. (*JEL* D83, R21, R31)

Housing markets are search markets. As a result, the typical description of a housing market includes not only price and trading volume, but also inventory available for sale, i.e., how many home sellers are searching for potential home buyers. In policy discussions, low inventory is often identified with a "housing shortage" in which housing demand outstrips housing supply. What is typically missing, however, is data on the number of potential buyers that are actually looking for a house in a given market. In addition, we usually do not observe information on the search behavior of these buyers.

Two key questions thus remain unanswered. First, is inventory really sufficient to summarize the state of the housing market? In other words, does the housing market resemble the labor market where low vacancy rates usually go along with high unemployment, as captured by a downward-sloping Beveridge curve? And second, how integrated are different housing market segments? Do home buyers usually focus their search activity on a few neighborhoods or do they consider entire metro areas? Answers to both questions are crucial to build quantitative models of housing search and make informed policy decisions.

*Piazzesi: Stanford and NBER (email: piazzesi@stanford.edu); Schneider: Stanford and NBER (email: schneidr@stanford.edu); Stroebel: New York University and NBER (email: johannes.stroebel@nyu.edu). John Leahy was the coeditor for this article. We thank Ed Glaeser, Bob Hall, Brian Higgins, Giuseppe Moscarini, Robert Shimer, Vincent Sterk, Silvana Tenreyro, four anonymous referees, and conference and seminar participants at the AEA meetings, Arizona State, Berkeley, Chicago, Harvard, LBS, Mannheim, Michigan, Minneapolis Fed, Minnesota, MIT, NBER Summer Institute, Northwestern, NYU, Philadelphia Fed, SITE, SED, Stanford, Tsinghua, UCLA, USC, and Yale. We also thank Trulia for providing data. This research is supported by an NSF grant.
[†]Go to https://doi.org/10.1257/aer.20141772 to visit the article page for additional materials and author disclosure statements.

This paper uses a novel dataset on the search behavior of home buyers to document stylized facts about housing search and to inform a quantitative housing search model. We show that the housing market in the San Francisco Bay Area is a collection of many small market segments that differ by geography and property characteristics. In every segment, there is local demand from narrow searchers who look for houses in only a handful of similar segments. At the same time, there are broader searchers who connect many segments to create integrated areas. For example, cities such as San Francisco and San Jose are broadly searched by many potential home buyers. We also document that the cross-sectional Beveridge curve depends on the level of aggregation. Across areas that do not share many common searchers, such as cities, the Beveridge curve slopes down as low-inventory cities experience high buyer search activity. In contrast, across housing market segments within cities, which share many common searchers, the Beveridge curve slopes up: segments with lower inventory draw interest from *fewer* searchers.

To understand these patterns and explore their implications, we build a model of housing search with three new features motivated by our empirical findings: (i) the housing market is a collection of many segments, (ii) agents differ in their search ranges, the set of segments they consider living in, and (iii) broad searchers flow into segments within their search range in proportion to segment inventory. The model explains how equilibrium housing market outcomes are shaped by the interaction of broad and narrow searchers.

In particular, different cities share few common searchers but differ in their popularity, the total number of agents who would consider living there. In more popular cities, higher demand means that a larger number of searchers quickly buys any house that comes on the market, so inventory is lower. As a result, the Beveridge curve slopes down across cities.

In contrast, segments within cities share many broad searchers but differ in their stability, the rate at which houses come on the market. In less stable segments, more supply means that higher inventory attracts more broad searchers who "crowd out" narrow searchers. This increases the total number of searchers per house. Less stable segments thus have more inventory and more search activity. As a result, the Beveridge curve slopes up across segments within cities, or more generally, across any set of segments with many common searchers.

The model also shows that buyer search data are essential for policy analysis. For example, to forecast market responses to new construction, it is important to avoid both too much and too little aggregation. Consider a researcher who is interested in the effect of construction on the inventory in a particular segment within a city. If she assumes incorrectly that the city is not at least partially integrated, she will miss the implications of competition between broad and narrow searchers. Indeed, building in a low-inventory segment may mostly attract broad searchers and thus neither increase inventory nor benefit narrow searchers in the segment. Similarly, researchers interested in metro area aggregates should avoid treating metro areas as fully integrated homogeneous markets. Indeed, we show that with partial integration, the overall change in metro area inventory depends on the popularity and stability of the exact segment where construction takes place.

We infer search ranges of potential home buyers from online housing search: on the popular real estate website trulia.com, home searchers can set an alert that triggers an email whenever a house with their desired characteristics comes on the market. We observe the search parameters in a large sample of such email alerts. Housing search occurs predominantly along three dimensions: geography, price, and house size as captured by the number of bathrooms (which are specified more often in searches than the number of bedrooms or the square footage). To relate search activity to other market activity, we divide the San Francisco Bay Area into 564 distinct housing market segments along the dimensions suggested by the observed search ranges. We then express the search ranges as subsets of the set of all segments, and measure search activity at the segment level by the number of searchers per house. We also measure the cross section of turnover and inventory at the segment level from deeds records, assessment data, and "for sale" listings.

Our search model assumes random matching as well as a fixed number of houses and agents. Houses are located in one of many segments, each with its own matching function. Moving shocks induce agents to sell their current house (at a cost) and search for another house. Heterogeneous agent types are identified by their search ranges: subsets of the set of all segments that they would consider living in, as in our data. While matching is random, agents are more likely to match in those segments within their search ranges where inventory is higher. This central assumption is directly supported by the patterns in our search data. We also show that it can be derived from more primitive assumptions in a variety of settings. Prices reflect the present value of housing services less a frictional discount due to search and transaction costs.

Our quantitative analysis proceeds in two steps. We begin by focusing on quantities. We show that the distribution of preferences, moving shocks, and a measure of matching frictions can be identified from cross-sectional moments of turnover, inventory, and search activity. This identification result is independent of the details of price bargaining and the matching function. We estimate that distribution and derive summary statistics of supply and demand conditions at the segment level. We define *stable* segments as those with less frequent moving shocks, and *popular* segments as those with a larger clientele (i.e., more individuals who are potentially interested in living there), where broader searchers count less toward popularity in any segment they cover. Our estimates of segment-level popularity are higher in areas with better schools, better restaurants, and better weather. As an overidentifying restriction test of our model, we show that population flows between segments implied by our estimates are consistent with observed moves in the data.

The estimated distribution of preferences explains why Beveridge curves differ by the level of aggregation. Bay Area cities are searched by fairly distinct clienteles and differ in their popularity. Variation in city popularity generates a downward-sloping housing Beveridge curve across cities: in more popular cities, any house that comes on the market is sold more quickly, leaving less inventory in equilibrium. More popular cities are also estimated to be more stable, which contributes to the downward slope of the Beveridge curve and helps explain why inventory and turnover comove positively across cities. At the same time, segments within cities are typically integrated by broad searchers who look to buy in the entire city, or at least all of its less expensive segments. These broad searchers are attracted to unstable segments where inventory regularly comes on the market. In those market segments, they crowd out any local narrow searchers, which generates a higher number of searchers per house.

In markets where broad and narrow searchers interact, variation in segment stability can therefore generate an upward-sloping Beveridge curve across segments.

To quantify the importance of the new mechanism of competition between broad and narrow searchers, we compare our estimated model to a misspecified benchmark that assumes all searchers narrowly target just one segment. This benchmark exercise follows the common approach in the literature of using data on turnover, inventory, and the time it takes to find a house to pin down parameters of a search model. In this benchmark without broad searchers, all observed search activity is attributed to local demand from narrow searchers. We show that this approach leads researchers comparing segments within a city to infer that unstable segments with high inventory are substantially more popular than they actually are, since the search activity by broad searchers attracted to the high inventory is indistinguishable from search activity by narrow searchers with a particular and targeted interest in that segment.

The second part of our quantitative analysis assumes bargaining over price as well as segment-specific Cobb-Douglas matching functions. We use the parameter estimates from the full model to study price formation and policy counterfactuals. We first infer frictional price discounts across segments, which capture the capitalized value of trading frictions faced by current and future buyers. These discounts are quantitatively large, between 5 percent and 35 percent of the frictionless house value (defined as the present discounted value of future housing services). Frictional discounts are larger in less stable segments, where houses transact more often. Frictional discounts are also larger in segments where houses take a long time to sell, for example because these segments do not attract many broad searchers. Quantitatively, the effect of search frictions on transaction prices is small compared to the effect of transaction costs.

In the final section of the paper, we explore how the response of housing markets to new construction depends on local clientele patterns. We contrast the construction of new housing in two neighborhoods that are similar in size and price, but differ in the share of broad searchers and hence their integration with the rest of the Bay Area. We find that new construction in urban San Francisco segments with many broad searchers affects segments across the entire city, but does not have a particularly large effect on inventory in the segments where construction actually takes place. In contrast, new construction in suburban segments close to the San Francisco city boundary, segments that mostly attract narrow searchers, increases inventory in those segments, with much smaller effects on nearby housing markets. Here, the number of narrow searchers who obtain housing surplus declines with new construction: the higher inventory makes it harder to sell houses, and increased competition from broad searchers attracted to the inventory makes it more difficult to buy. The effect of construction on aggregate Bay Area inventory also depends on the segment in which the construction takes place.

We conclude that information on clientele patterns is essential for housing policy. Too much aggregation leads to bias in predicting the effects of construction on both local and aggregate inventory. Indeed, a typical calibration to aggregate metro area moments will infer an elasticity of housing demand as if there were only broad searchers. Most concrete zoning proposals, however, focus on building in a particular area. If the metro area is not fully integrated, local clientele patterns can lead to larger or smaller responses of equilibrium inventory than one would predict using an aggregate model. Our examples suggest that these effects can be quantitatively large.

Relatedly, our results on the Beveridge curve call for a cautious and selective use of low inventory as a signal of a housing shortage. Building houses in cities with low inventory does indeed address a housing shortage: it targets cities where local excess demand is high and all new houses satisfy local demand. This reduces the time it takes households that are interested in this city to find a house, and increases steady-state inventory. In contrast, building houses in low-inventory neighborhoods within a city selects segments with relatively low local excess demand. Instead, the additional inventory mostly attracts more broad searchers who have no specific preference for the location. The inflow of broad searchers who now compete with the narrow searchers for the additional inventory reduces the effect that the new construction has on the time it takes these narrow households to find a home.

Our paper contributes to a growing body of empirical work that analyzes housing market activity in the cross section. The typical approach is to sort houses within geographic units (such as cities) based on similarity along property characteristics, and to then compare measures of market activity such as turnover and time on market across these segments (see Goodman and Thibodeau 1998, Leishman 2001, Islam and Asami 2009). There is, however, little direct evidence on housing search behavior. An exception is Genesove and Han (2012), which builds a time series of search activity at the city level from survey data on buyers' house hunting experiences. Our paper offers a new source of demand-side information and uses the distribution of online searchers' criteria to define segments for which market activity is then measured.[1] Our approach emphasizes the heterogeneity of market and search activity within cities and across potential buyers.[2]

We build on a large literature on housing search models (see Han and Strange 2015 for a review). Recent studies have turned to quantitative evaluations using microdata (see Díaz and Jerez 2013; Head, Lloyd-Ellis, and Sun 2014; Ngai and Tenreyro 2014; Guren and McQuade 2014; Anenberg and Bayer 2014; Halket and Pignatti 2015; Burnside, Eichenbaum, and Rebelo 2016; Guren 2018). Most of these authors are interested in how search modifies the time series dynamics of house prices and market activity in homogeneous housing markets. In contrast, our focus is on steady states of markets with rich cross sections of houses and buyers. Our theoretical model is based on earlier random matching models such as Wheaton (1990), Krainer (2001), Albrecht et al. (2007), and Piazzesi and Schneider (2009); the new element is that we allow for matching by heterogeneous buyers across multiple interconnected segments.

[1] Our paper therefore contributes to an emerging literature that shows how the increasing availability of data from online services such as eBay, Facebook, Trulia, and others allows researchers to overcome important measurement challenges across the social sciences (see, for example, Einav et al. 2015; Bailey et al. 2017, 2018a, b, 2019a, b).

[2] As such, our paper also contributes to the literature that considers house valuation with heterogeneous regions and buyers. For example, Poterba (1991) considers the role of demographics for prices; Bayer, Ferreira, and McMillan (2007) looks at school quality; Giglio et al. (2015) explores the effects of climate change risk; and Guerrieri, Hartley, and Hurst (2013) studies the effects of gentrification. Stroebel (2016) and Kurlat and Stroebel (2015) investigate the market impact of asymmetric information about property and neighborhood characteristics, respectively. Landvoigt, Piazzesi, and Schneider (2015) studies the effect of credit constraints on prices in an assignment model with many quality segments. They consider competitive equilibria of a model with homogeneous preferences. In contrast, this paper emphasizes frictional discounts due to search and transaction costs, as well as heterogeneity in search preferences within a metro area. Davis and Van Nieuwerburgh (2015) and Piazzesi and Schneider (2017) survey the broader literature on business cycles, asset prices, and housing, including studies that do not rely on search frictions.

The paper in the labor search literature that is closest to our work is Manning and Petrongolo (2017), which estimates a search and matching model for local labor markets. They divide space into small areas that searchers can search jointly, thus generating spillover effects across regions. Their demand-side data consist of job seekers' home addresses and the vacancies they apply to. They infer a distribution of preferences that includes a parameter for how fast utility declines with commuting time. Our detailed search data allow us to put less structure on utility, and to use both spatial and quality information to define the commodity space.

The rest of the paper is structured as follows. Section I describes our housing search data, and presents key patterns of housing search behavior. Section II establishes stylized facts on housing market and search activity at the segment level, exploring how the slope of the housing Beveridge curve varies with the level of aggregation. Section III presents a reduced-form model of a single segment that highlights the key economic forces arising from the interaction of broad and narrow searchers in housing markets. Section IV estimates a fully fledged housing search model with many segments that quantifies the importance of these forces. Section V uses estimates from this model to infer frictional discounts and to explore the response of prices and quantities in different market segments to local housing market shocks. The final section concludes by discussing the importance of understanding the segmentation of search clienteles across a number of other important search markets, such as over-the-counter financial markets, labor markets, and dating markets.

I. Understanding Housing Search Behavior

We document housing search behavior in the San Francisco Bay Area using email alerts set on the popular real estate website trulia.com. The San Francisco Bay Area is a major urban agglomeration in Northern California that includes San Francisco, San Jose, Oakland, as well as a number of other cities. We analyze data from Alameda, Contra Costa, Marin, San Benito, San Francisco, San Mateo, and Santa Clara counties. In the 2010 Census, these counties had a population of about 6 million people living in 2.2 million housing units. This section first describes the email alert data and then highlights the key findings from online Appendix Section A, where we provide a detailed analysis of the raw search patterns.

A. *Search Data from Email Alerts*

Visitors to trulia.com can set alerts that trigger regular emails when houses with certain characteristics come on the market. Every alert must specify interest in houses that are "For sale," "For rent," or "Recently sold." Each alert must also specify a location of interest that allows for a list of zip codes, neighborhoods, or cities. Neighborhoods are geographic units commonly listed on realtor maps that are often aligned with zip codes. When users fill out the form, an auto-complete function suggests names of neighborhoods or cities.

The second row in the form provides the option of specifying property characteristics beyond geography. Price ranges may be set by providing a lower bound, an upper bound, or both. For bedrooms, bathrooms, and house size, there is the option

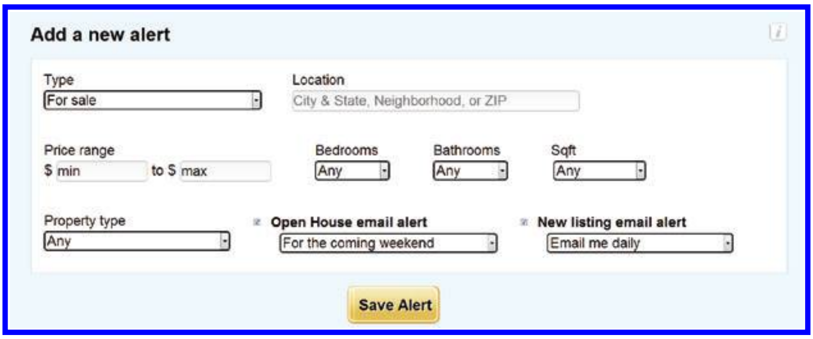

Figure 1. Setting Email Alerts on Trulia.com

Note: Figure shows the web form for setting email alerts on trulia.com.

to set a lower bound or an upper bound. In the third row, "Property type" allows narrowing the search to "Single family home," "Condo," and several smaller categories. The remaining fields govern how emails are processed: for the "New listing email alerts" relevant to our study, the options are to receive a daily or weekly email.[3]

We observe 40,525 "For sale" email alerts set for Bay Area properties between March 2006 and April 2012. Those alerts were set by 23,597 unique home searchers, identified by their (scrambled) email addresses. Almost 70 percent of searchers set only one alert, and more than 90 percent of searchers set three or fewer alerts. Since we are interested in search ranges rather than individual alerts, we pool alerts set by the same searcher, as described in online Appendix Section A.

Representativeness of Search Behavior.—We do not observe demographic information on the home searchers in our sample. Thus, we cannot provide direct evidence that searchers on trulia.com are representative of the overall pool of home searchers. However, surveys conducted by the National Association of Realtors (2013) during our sample period suggest that the internet is the most important tool in the modern home-buying process. Indeed, over 90 percent of home buyers rely on the internet in their search. The fraction of people who deemed real estate websites "very important" as a source of information was 76 percent, substantially larger than the 68 percent who found real estate agents "very important."

Internet use for home search is also not concentrated among younger or richer buyers: 86 percent of home buyers between the ages of 45 and 65 go online to search for a home. The median age of home buyers using the internet is 42, and their median income is $83,700 (National Association of Realtors 2011). This is only slightly younger than the median of all home buyers (which is 45), and only slightly wealthier (the median income of all home buyers is $80,900). These statistics suggest that we can learn from online home search about overall home search behavior. Moreover, trulia.com, with approximately 24 million unique monthly visitors during our sample period (71 percent of whom report planning to purchase in the

[3] Trulia also provides a second way for potential home buyers to set an email alert. After looking at results from regular searches on their website, generally along the same dimensions as those in Figure 1, users can press a single button: "Send me an email whenever houses with these characteristics come on the market."

next 6 months), has similar user demographics to those of the overall online home search audience (Trulia 2013).

B. *Dimensions of Housing Search*

To compare the geographic dimension of individuals' search ranges, we express them in terms of the zip codes that are covered by the pooled email alerts. We observe wide heterogeneity in the geographic breadth considered by various home searchers. While 25 percent of searchers are narrowly interested in a single zip code, among those individuals who select more than one zip code, the 10–90 percentile range of the maximum geographic distance between zip codes selected by the same searcher is 2.3 miles to 21.1 miles. Online Appendix Section A.2 provides further details.

Roughly two-thirds of the email alerts include search parameters in addition to geography. The other fields that are specified regularly are listing price (two-thirds of all alerts) and the number of bathrooms (one-third of all alerts). On the price dimension, among those searchers who set both an upper and a lower bound, the tenth percentile selects a price range of \$100k, the median selects a price range of \$300k, and the ninetieth percentile selects a price range of \$1.1m. Among the same searchers, the median person specifies a price range of ± 27 percent around the midpoint of the range. At the tenth percentile of the distribution, this figure is ± 12.5 percent around the midpoint, and at the ninetieth percentile it is ± 58 percent. Among those searchers who specify the number of bathrooms, almost all select a lower bound of "2."

We want to develop a model that captures the heterogeneity of these search ranges. One possible approach to summarize their geographic dimensions would be to use contiguous and/or circular subsets of a plane. This approach does not work well for the Bay Area with its complicated topology. Moreover, many searchers look for houses in zip codes that are not necessarily adjacent to each other. Our approach in the next section is therefore to define a discrete grid of market segments using zip codes as the basic geographic unit, which we further subdivide along quality and size dimensions. A search range can then be represented as a subset of market segments, allowing us to accommodate the observed noncontiguous and noncircular search patterns.

II. Segment-Level Housing Search and Market Activity

We now describe how we divide the San Francisco Bay Area into a finite number of housing market segments, motivated by the search ranges inferred from our email alert data. We then establish stylized facts on market and search activity at the segment level.

The finest partition of the Bay Area housing stock into different market segments that could be motivated by our search data is obtained by the join of all search ranges in our sample. The preferences of each searcher could then be expressed exactly as a subset of these segments. However, the problem with this approach is sample size: the number of houses per segment would be too small to accurately measure moments such as inventory and time on market.

Our approach, therefore, is to get as close as possible to the finest partition, but subject to the constraint that segments must be sufficiently large in terms of volume and housing stock. We start with zip codes as the level of geography, and then subdivide zip codes along price and size boundaries that are common in email alerts that cover a particular zip code. We then merge small but similar segments within a zip code until all remaining segments have at least 1,500 housing units, or no further merges are possible. This process leads us to a final set $\mathcal{H}$ of 564 segments that are sufficiently large to accurately measure housing market activity. These segments contain houses within a zip code that are of similar quality (based on price) and size (based on the number of bathrooms), with cutoffs that are close to cutoffs regularly used in email alerts. Online Appendix Section B.1 provides details on the algorithm.

We then express each search range as a subset of these segments. Here we start from the raw search ranges, specified along the dimensions geography, quality, and size; we ignore the other dimensions that are rarely specified. We then determine which segments are (approximately) covered by each range, using an algorithm and cleaning procedure described in online Appendix Section B.2. This produces a set Θ of 4,956 distinct search ranges that is each represented as a subset of $\mathcal{H}$.

A. *Segment-Level Housing Market Activity*

To measure housing market activity at the segment level, we combine three main datasets. We start with the universe of ownership-changing deeds in the San Francisco Bay Area between 1994 and April 2012. From the deeds data, we obtain the property address, transaction date, transaction price, and the type of deed (e.g., Intra-Family Transfer Deed, Warranty Deed). We use the type of deed to identify arms-length transactions (see online Appendix Section B.3 for details). We combine these transaction deeds with the universe of tax assessment records for the year 2009. This dataset includes information on property characteristics such as construction year, size, and the number of bedrooms and bathrooms. Finally, we use data on all property listings on trulia.com between October 2005 and December 2011. The key variables from this dataset are listing date, listing price, and the listing address. The latter can be used to match listing data to deeds data. We can then construct a measure of time on market for each property that eventually sells, as well as the inventory that is for sale in a market segment at each point in time.

Throughout our analysis, we pool observations for the period 2008–2011, a time period for which we observe information on both housing search and housing market activity. The goal of this paper is to understand the cross section of market activity. Pooling observations across years helps us achieve a finer description of cross-sectional heterogeneity by ensuring that there are sufficiently many observations to measure market activity in segments with few listings and low housing turnover. In online Appendix Section C, we show that segment-level market and search activity are quite stable over time within our sample period. To make prices comparable across years, we convert all prices to 2010 dollars using zip code-level repeat sales price indices.

Segment-Level Facts: Notation.—We next present segment-level facts about market and search activity. These facts reveal a number of interesting patterns that motivate the subsequent quantitative exercise. The following notation, summarized in Table 1, is useful to organize facts reported at the segment level. As before, $\mathcal{H}$

Table 1—Main Notation in the Paper

Symbol	Definition	Additional details
Segment-level notation		
$\mathcal{H}$	Set of all segments h	$h \in \mathcal{H}$
$H(h)$	Housing stock in h	Normalization: $\sum_{h\in\mathcal{H}} H(h) = 1$
$m(h)$	Transactions in h	
$V(h)$	Monthly turnover rate in h	$V(h) = m(h)/H(h)$
$T(h)$	Average time on market in h	Measured in months
$L(h)$	Inventory in h	Steady state: $L(h) \coloneqq T(h)m(h)$
$I(h)$	Inventory share in h	$I(h) = L(h)/H(h)$
$v(h)$	Flow of housing services in h	
$p(h)$	Transaction price in h	See online Appendix equation (E.2)
$\sigma(h)$	Weighted searchers per house in h	$\sigma(h) = \frac{1}{H(h)} \sum_{\theta\in\tilde{\Theta}(h)} \frac{\tilde{B}(\theta)}{\bar{B}} \frac{H(h)}{\tilde{H}(\theta)}$
$\tilde{\Theta}(h)$	Clientele of h	$\tilde{\Theta}(h) = \left\{\theta \in \Theta : h \in \tilde{\mathcal{H}}(\theta)\right\}$
$B(h)$	Buyers in h	$B(h) = \sum_{\theta\in\tilde{\Theta}(h)} \tilde{B}(\theta) \frac{L(h)}{\tilde{L}(\theta)}$
$\tilde{m}(B(h), L(h), h)$	Matching function in h	
$\eta(h)$	Instability of h	
$\alpha(h)$	House finding rate in h	$\alpha(h) = m(h)/B(h)$
$\pi(h)$	Popularity of h	$\pi(h) \coloneqq \frac{1}{H(h)} \sum_{\theta\in\tilde{\Theta}(h)} \mu(\theta) \frac{H(h)}{\tilde{H}(\theta)}$
Searcher-level notation		
Θ	Set of search ranges θ	$\theta \in \Theta$
$\mu(\theta)$	Measure of θ	$\sum_{\theta\in\Theta} \mu(\theta) = \bar{\mu} > 1$
$\tilde{\mathcal{H}}(\theta)$	Segments scanned by θ	$\tilde{\mathcal{H}}(\theta) \subset \mathcal{H}$
$\tilde{H}(\theta)$	Housing stock of interest to θ	$\tilde{H}(\theta) = \sum_{h\in\tilde{\mathcal{H}}(\theta)} H(h)$
$\tilde{L}(\theta)$	Inventory considered by θ	$\tilde{L}(\theta) = \sum_{h\in\tilde{\mathcal{H}}(\theta)} L(h)$
$\tilde{B}(\theta)$	Buyers of type θ	$\sum_{\theta\in\Theta} \tilde{B}(\theta) = \bar{B}$
Other notation		
c	Proportional transaction cost	

denotes the set of all 564 segments. The housing stock in segment h is given by $H(h)$. We normalize the total Bay Area housing stock to a unit mass: $\sum_{h\in\mathcal{H}} H(h) = 1$.

The average monthly turnover rate $V(h)$ in segment h is defined as the number of arms-length transactions in that month $m(h)$ divided by the housing stock $H(h)$. The mean time on market $T(h)$ in segment h is defined as months between listing and sales date, less one month for the typical escrow period. Our measure of inventory in segment h is $L(h) \coloneqq T(h)m(h)$.[4] We also define the inventory share $I(h) = L(h)/H(h)$ as the share of the housing stock that is for sale.

Every search range in our sample is a subset of the set of all segments $\mathcal{H}$. We index the ranges by $\theta \in \Theta$ and refer to the set Θ as the set of "searcher types."

[4] This measure of inventory conditions on houses that are eventually sold, since time on market $T(h)$ is based on actual sales. Alternatively, one could construct measures of inventory directly from listings data. The resulting data series are noisy because of the incomplete coverage of Trulia listings data, and the need to make assumptions on when the few listings that do not sell are removed. We discuss the trade-offs involved in the choice of how to measure inventory in online Appendix Section B.3.

A searcher of type θ scans inventory in the set of segments $\tilde{\mathcal{H}}(\theta) \subset \mathcal{H}$. The total housing stock that is of interest to searcher θ is

$$\tilde{H}(\theta) = \sum_{h \in \tilde{\mathcal{H}}(\theta)} H(h).$$

Similarly, we define the total inventory considered by searcher θ as $\tilde{L}(\theta) = \sum_{h \in \tilde{\mathcal{H}}(\theta)} L(h)$, which is the sum over all inventory for sale in segments in θ's search range $\tilde{\mathcal{H}}(\theta)$. The *clientele* of segment h consists of all searchers who consider segment h as part of their search range, that is,

$$\tilde{\Theta}(h) = \left\{\theta \in \Theta : h \in \tilde{\mathcal{H}}(\theta)\right\}. \tag{1}$$

The pattern of clienteles reflects the interconnectedness of segments. One polar case is a perfectly segmented market, in which $|\mathcal{H}|$ types have search ranges that each consist of a single segment, and each segment has a homogeneous clientele of one type who searches only in that segment. Another polar case is a perfectly integrated market, where a single type searches across all segments and all clienteles are identical and contain only that type. More generally, clienteles are heterogeneous and may consist of distinct types with only partially overlapping search ranges.

Let $\tilde{B}(\theta)$ denote the number of buyers with search range θ. The total number of buyers is $\sum_{\theta \in \Theta} \tilde{B}(\theta) = \bar{B}$. The distribution of searchers interested in segment h is then obtained by integrating the distribution $\tilde{B}(\theta)/\bar{B}$ over $\tilde{\Theta}(h)$. We want a summary statistic of search activity that is comparable across segments that differ in the number of broad and narrow searchers. We compute a measure of *weighted searchers per house* in segment h that weights searchers according to their search breadths:

$$\sigma(h) = \frac{1}{H(h)} \sum_{\theta \in \tilde{\Theta}(h)} \frac{\tilde{B}(\theta)}{\bar{B}} \frac{H(h)}{\tilde{H}(\theta)} = \sum_{\theta \in \tilde{\Theta}(h)} \frac{\tilde{B}(\theta)}{\bar{B}\tilde{H}(\theta)}. \tag{2}$$

Weighting the contribution of each searcher type by $H(h)/\tilde{H}(\theta)$ makes broader searchers count less toward search activity in segment h. If every searcher was looking at only one segment, then $\sigma(h)$ would simply reflect the relative number of searchers per house in h, since in that case the housing stock $\tilde{H}(\theta)$ of a searcher θ interested in h would equal $H(h)$. More generally, our measure of search activity $\sigma(h)$ is an index which is normalized to one for the entire Bay Area.[5] So far, all summary statistics have been defined at the segment level. We are also interested in how market and search activity vary at different levels of aggregation. Since V, I, and σ are all defined as ratios relative to the housing stock, aggregation uses the housing stock as weights. For example, the turnover rate over some subset $G \subset \mathcal{H}$, such as a zip code or city, is computed as

$$\frac{\sum_{h \in G} H(h) V(h)}{\sum_{h \in G} H(h)}.$$

[5] The sum over all $\sigma(h)$s in the Bay Area, weighted by housing stock, is

$$\sum_{h \in \mathcal{H}} \frac{H(h)}{1} \sigma(h) = \sum_{h \in \mathcal{H}} \sum_{\theta \in \tilde{\Theta}(h)} \frac{\tilde{B}(\theta)}{\bar{B}} \frac{H(h)}{\tilde{H}(\theta)} = \sum_{\theta \in \Theta} \sum_{h \in \tilde{\mathcal{H}}(\theta)} \frac{\tilde{B}(\theta)}{\bar{B}} \frac{H(h)}{\tilde{H}(\theta)} = \sum_{\theta \in \Theta} \frac{\tilde{B}(\theta)}{\bar{B}\tilde{H}(\theta)} \sum_{h \in \tilde{\mathcal{H}}(\theta)} H(h) = \sum_{\theta \in \Theta} \frac{\tilde{B}(\theta)}{\bar{B}} = 1.$$

Table 2—Summary Statistics of Market and Search Activity

	Inventory share I (in percent)	Turnover rate V (in percent)	Search activity σ	Mean price (in k$)	Housing stock
Bay area	1.14	0.24	1.00	650	2,216,021
Across-segment					
Min	0.00	0.00	0.05	72	1,221
P25	0.61	0.16	0.51	306	2,333
P50	0.94	0.21	0.81	519	3,298
P75	1.51	0.29	1.31	790	4,858
Max	6.66	0.97	7.05	2,491	13,167

Notes: Table shows summary statistics for market and search activity, both for the entire San Francisco Bay Area as well as across the 564 housing market segments. We consider inventory share I, turnover rate V, search activity σ as defined in equation (2), mean price, and total housing stock.

Segment-Level Facts: Summary Statistics.—Table 2 presents summary statistics on market and search activity for the San Francisco Bay Area as a whole, as well as by segment. The housing market is relatively illiquid, which is consistent with our sample period 2008–2011 covering the housing bust when housing transactions were unusually low. On average, only 1.14 percent of the Bay Area housing stock is for sale at any point in time. The average monthly turnover rate is 0.24 percent, so that the typical house turns over once every $100/(0.24 \times 12) = 35$ years. The cross-sectional variation in market activity at the segment level is substantial. For example, the seventy-fifth percentile of inventory share is 1.51 percent, over twice as much as the 0.61 percent inventory share at the twenty-fifth percentile.

The distribution of $\sigma(h)$, our measure of search activity, is positively skewed across segments: most segments have less than one weighted searcher per house. The minimum of 0.05 is achieved by a segment in San Jose, which is only considered by a few broad searchers. Other segments have substantially more search activity, all the way to a maximum of 7.05 in a segment in central San Francisco, which attracts many narrow searchers in addition to broad searchers.

Segment-Level Housing Market Activity: Within and across Submarket Correlation.—Table 3 reports cross-sectional correlations of observables for "submarkets" at different levels of aggregation. These submarkets are the 564 segments, the 191 zip codes, and the 96 cities in our data. The left panel shows volatilities and correlation coefficients *across* submarkets. The right panel considers segment-level variation *within* submarkets.

A comparison of volatilities shows substantial variation across segments within the same zip code. Indeed, the zip code-level movements account for only 46 percent, 47 percent, and 65 percent of the across-segment variance in inventory share I, turnover rate V, and search activity σ, respectively.

The comovement of search activity (σ) and market activity $(V$ and $I)$ depends crucially on the level of aggregation. While it is close to zero at the segment level, it turns negative when analyzed across zip codes, and even more negative when analyzed across cities. In addition, more expensive zip codes and cities are searched more. In contrast, search activity comoves positively with inventory and turnover across segments within zip codes or cities. Within cities and zip codes, more expensive segments are searched less.

Table 3—Cross-Sectional Variation in Market and Search Activity

	Variation across cities					Average variation within cities			
	I	V	σ	$\log(p)$		I	V	σ	$\log(p)$
Volatility	0.42	0.07	0.49	0.46	Volatility	0.54	0.09	0.40	0.42
Correlation	1	0.93	−0.42	−0.74	Correlation	1	0.88	0.51	−0.68
		1	−0.34	−0.60			1	0.42	−0.52
			1	0.53				1	−0.51
				1					1
	Variation across zip codes					Average variation within zip codes			
Volatility	0.50	0.08	0.65	0.50	Volatility	0.58	0.09	0.55	0.41
Correlation	1	0.92	−0.18	−0.68	Correlation	1	0.83	0.58	−0.76
		1	−0.15	−0.53			1	0.41	−0.61
			1	0.41				1	−0.55
				1					1
	Variation across segments								
Volatility	0.74	0.12	0.81	0.64					
Correlation	1	0.93	0.03	−0.63					
		1	0.01	−0.49					
			1	0.08					
				1					

Notes: Table shows the cross-sectional variation in market and search activity at different levels of aggregation: 564 segments, 191 zip codes, and 96 cities. The left panel presents statistics across these units, the right panel presents statistics across segments within these units. We present inventory share I, turnover rate V, search activity σ, and log(price). We present both volatilities (standard deviations) as well as correlations within and across units.

The relationship between inventory and measures of search activity is reminiscent of the "Beveridge curve" in studies of the labor market. The Beveridge curve is a relationship between vacant job positions and unemployed workers searching for jobs, usually presented in the *time series*. Here we have a relationship between vacant homes and individuals searching for homes, but presented in the *cross section*. The Beveridge curve is *downward-sloping* across broad units of aggregation such as cities, while it is on average *upward-sloping* across small segments within broad units. Panel A of Figure 2 shows the Beveridge curve relationship across Bay Area cities, and panel C across all segments within the city of San Francisco. The San Francisco patterns are not unusual. Indeed, the Beveridge curve is upward-sloping in 64 out of the 74 cities with at least two segments, representing 84 percent of the total Bay Area housing stock. This fact is also not primarily driven by small cities: the within-city across-segment Beveridge curve slopes up for 23 of the largest 25 cities by housing stock, with an average correlation of 0.42, and a twenty-fifth percentile correlation of 0.21.

For our market activity indicators and prices, the nature of covariation across and within cities is very similar, both qualitatively and quantitatively: the inventory share and the turnover rate comove positively, and both are negatively correlated with price.[6] To illustrate this, panel B of Figure 2 shows the positive correlation between inventory shares and turnover rates across cities in the San Francisco Bay

[6]The high-turnover low-price segments could correspond to segments with many starter homes (see Ortalo-Magné and Rady 2006).

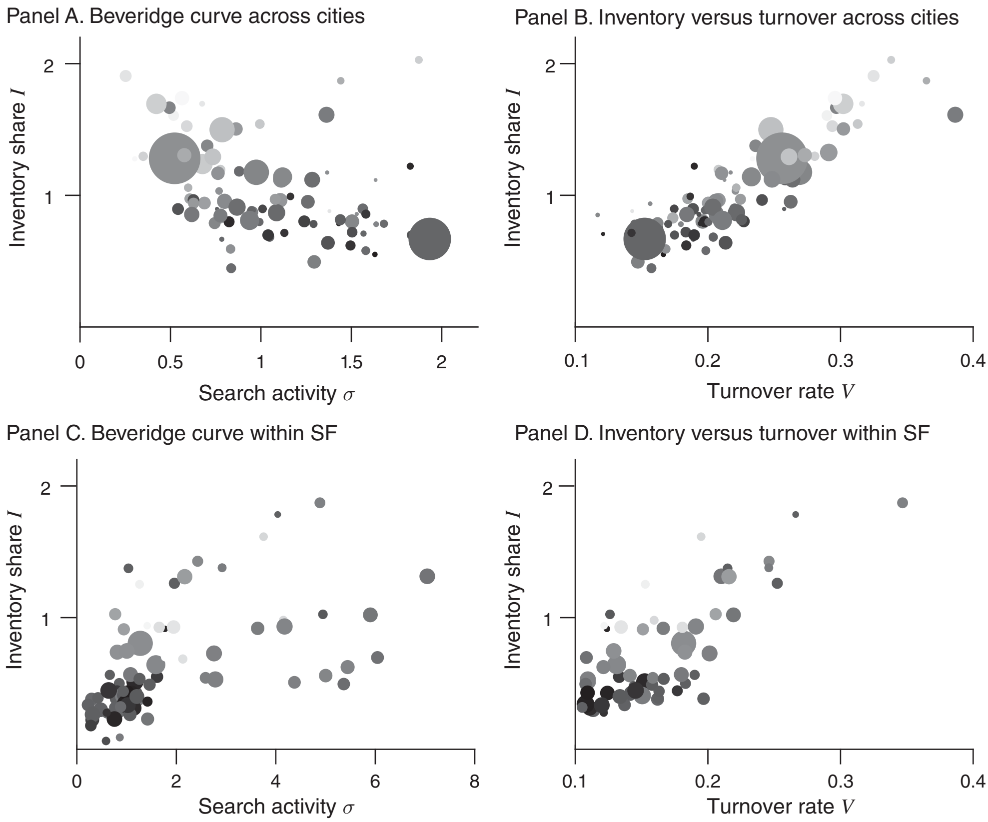

FIGURE 2. INVENTORY, SEARCH ACTIVITY, AND VOLUME

Notes: Figure shows the inventory share I, search activity σ, and the turnover rate V, both across and within cities. Panel A shows inventory shares and search activity (the Beveridge curve) across cities. Panel B shows inventory shares and turnover rates across cities. Panel C shows the Beveridge curve across segments within the city of San Francisco. Panel D shows inventory shares and turnover rates across segments within San Francisco. The size of the dots reflects the size of the house stock (from small to large). The color of the dots reflects housing prices from cheap (light) to expensive (dark).

Area. Panel D shows the positive correlation between inventory shares and turnover rates across segments within the city of San Francisco.

Segment-Level Facts: Search-Breadth.—The measure $\sigma(h)$ reflects the average search activity in a segment, but it does not tell us whether that activity is due to narrow searchers or due to broader searchers who provide connections to other segments. Indeed, the same $\sigma(h)$ could arise when there are a few narrow searchers who fully target their search effort to a given segment, or when there are relatively more broad searchers whose search effort in a given segment is diluted because they also consider other segments. To summarize interconnectedness, we compare segments in terms of the inventory scanned by their median client. Panel A of Figure 3 plots the share of inventory in segment h in total Bay Area inventory against the share of Bay Area inventory scanned by the median client of segment h. Every dot represents

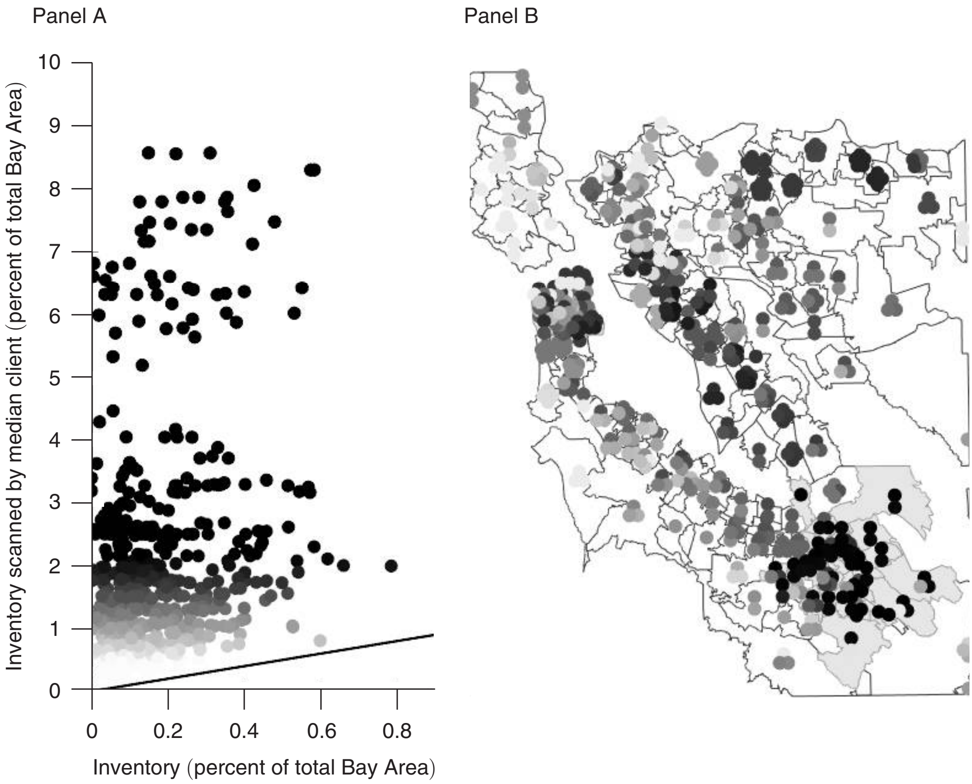

Figure 3. Scanned Inventory

Notes: Figure shows scanned inventory by median searcher in each segment, presented as a share of total Bay Area inventory. In both panels, each dot corresponds to one segment. In panel A, the horizontal axis shows the inventory in that segment as a fraction of total Bay Area inventory. The vertical axis shows the fraction of total Bay Area inventory scanned by the median searcher in that segment. The straight line is the 45-degree line. Panel B shows the geographic distribution of these segments. Dots for segments within the same zip code are arranged clockwise by price with the lowest priced segment at noon.

a segment, and colors reflect the value on the vertical axis so that segments can be recognized in the map in panel B.

If the Bay Area was perfectly segmented, then any given segment would only be searched by individuals who scan that particular segment. As a result, all dots would line up along the 45-degree line (which is the straight line in panel A of Figure 3). At the opposite extreme, if the Bay Area were perfectly integrated, then every client of every segment would scan all houses, so all dots should line up along a horizontal line at 100 percent of total inventory (located north of our current figure). The reality is in the middle: the median searcher in a segment scans multiple times more inventory than is available in the segment itself, but far less than the Bay Area total.

Online Appendix Section B.4 provides summary statistics on this measure of search breadth. The median searcher in the average segment scans 2.1 percent of the total Bay Area inventory; in segments at the twenty-fifth percentile of the distribution, the median searcher scans 0.9 percent of Bay Area inventory, while in segments at the seventy-fifth percentile of the distribution, the median searcher scans 2.5 percent of Bay Area inventory. Remarkably, the average within-segment interquartile range of inventory scanned by different searchers looking in the same segment is similarly large, at 1.8 percent of the Bay Area inventory. This highlights that while segments differ in the average search breadth of their clientele, most segments feature competition between narrow and broad searchers. The interaction between these broad and narrow searchers will be a central force in our model.

Areas with a large common clientele appear in panel A as near-horizontal clusters: if any subset of segments were perfectly integrated but not connected to other segments, then it would form a horizontal line at the level of its aggregate inventory. This effect is visible for the top cluster of dark black dots. The map in panel B shows that those dots represent cheaper segments in the city of San Jose, which is marked gray. More generally, clusters of dots with high scanned inventory correspond to cheap urban areas, where broad search appears to be more common.

B. *Search Intensity within Search Range*

How do searchers choose among the inventory within their search range? Are searchers equally interested in all properties, or do they prefer properties in some segments in their search range to those in others? To address this question, we exploit data on property views by individual home searchers on trulia.com. After defining a search range on trulia.com, a user is presented with a list of properties that are included in that search range (see online Appendix Figure A.7 for a screenshot of the interface). This list provides basic information on each property, such as its location, the listing price, a picture, and the first lines of a description of the property. Home searchers then actively click on those houses that attract their particular interest to view additional property information and potentially contact the realtor representing the seller.

We have obtained data on such detailed property views from trulia.com for a random subset of users visiting the site in April 2012. These data contain the set of listings viewed within a "session," defined as all views by the same user (represented by their IP address) within one day. We interpret viewing a property's listing details as an expression of particular interest in that property. In online Appendix Section A.5, we analyze how this signal of particular interest is distributed across listings in the various segments searched by an individual. We find that the rate at which particular interest is expressed for properties in different segments is directly related to the share of total inventory made up by those segments in the individuals' overall search ranges. This finding suggests that, conditional on the search range, the probability of finding a favorite house in any one particular segment is proportional to the inventory in that segment. This observation motivates one of our key modeling assumptions in the next section.

III. A Stylized Model of a Single Segment

In this section, we consider a simple reduced-form model of a single segment. We show how the cross section of observables introduced in Section IIA (turnover, inventory, and search activity) is shaped by supply and demand forces. The equations of the reduced-form model describe flows of buyers and houses in a particular segment. They hold in steady-state equilibrium for a large class of continuous-time search models. We specify and estimate one such model in Section IV. The purpose of the discussion below is to illustrate the key mechanisms that shape the housing Beveridge curve independently of model details such as price formation and how broad searchers select houses within their search ranges. Most importantly, we show that the interaction of broad and narrow searchers explains the key stylized fact about the Beveridge curve within and across cities from Figure 2. Moreover, we show how policy analysis at the wrong level of aggregation can lead to misleading conclusions.

Consider a segment with mass H of houses. If L houses are listed for sale and B agents are looking to buy, transactions occur at the rate $m(B,L)$, where the matching function m is increasing in both arguments and exhibits constant returns to scale. Agents own at most one house. The $H - L = H(1 - I)$ homeowners who do not already list their house for sale receive moving shocks at a constant rate η. Upon receiving a shock, they list their house, but stay in it until they sell; only then do they look for a new house. Individual agents thus cycle across three states: owning but not listing, owning and listing, and looking for a new house. In steady state, the number of agents in each state is constant. The share of houses coming on the market must thus be the same as the turnover rate at which houses are sold, or $\eta(1 - I) = V$. We now characterize the distribution of agents under different assumptions about buyer behavior.[7]

Narrow Searchers and the Cross Section of Cities.—Suppose first that all agents are narrow types who only want to live in the segment under study. If the total number of narrow-type agents is $N > H$, the number of agents looking for a house is $B^N = N - H$. All houses are owned by $N - B^N$ narrow types. In steady state, the number of houses those owners put up for sale must equal the number of houses sold:

$$\eta(1 - I)(N - B^N) = m\left(B^N, I(N - B^N)\right). \tag{3}$$

The equation implies an equilibrium relationship between the number of narrow searchers and the inventory share: if a larger share of houses is for sale, then it is easier for narrow types to find a house so the number of narrow searchers is smaller.

The model with narrow types is useful to interpret our scatter plots on the cross section of cities. Panels A and B of Figure 4 are designed to explain the variation in the scatter plots in panels A and B of Figure 2. Panel A relates inventory share and search activity $\sigma = B^N/H$ for two segments of the same size H, indicated by different line styles. For each segment, there is a black downward-sloping curve described by (3) as well as well as a gray vertical bar to indicate the exogenous number of narrow searchers. Solid lines describe a baseline segment and the baseline equilibrium inventory share and search activity are intersections marked by circles. Panel B relates inventory share and turnover: the gray downward-sloping curve is the condition $\eta(1 - I) = V$: since houses come on the market at a constant rate,

[7] Formally, an individual agent's state evolves according to a stationary Markov chain. In steady state, the ergodic distribution of this chain is the same as the distribution of agents in the population.

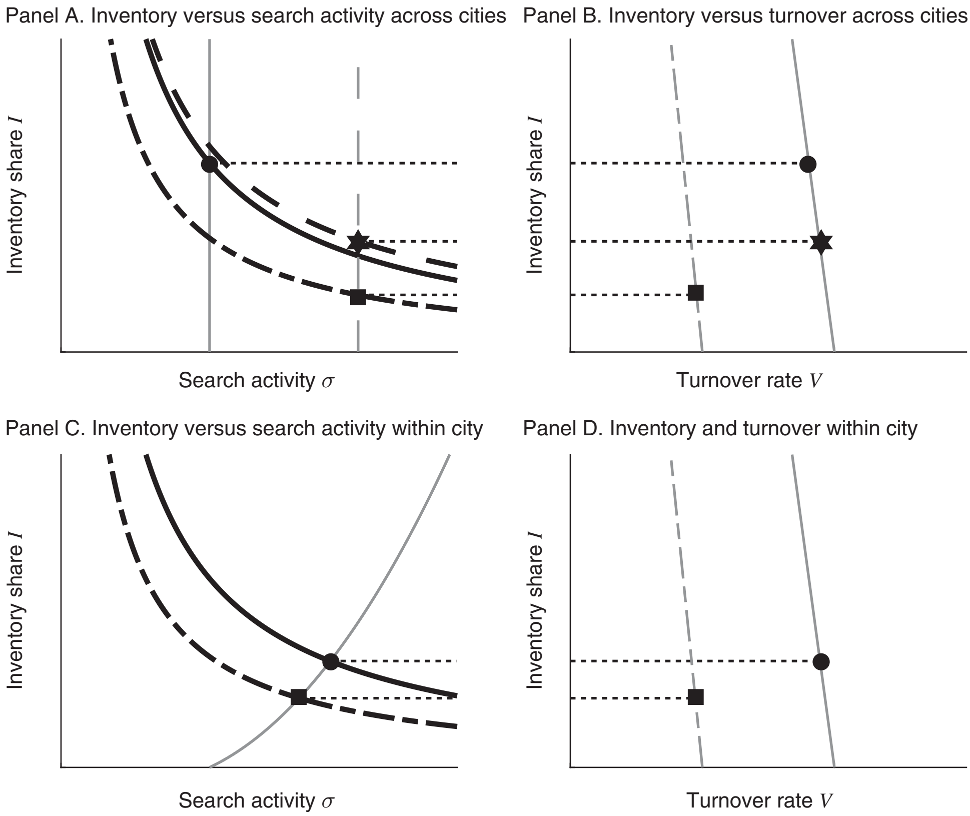

Figure 4. Inventory, Search Activity, and Volume in Stylized Single-Segment Model

Notes: Figure shows inventory I, search activity σ, and turnover rate V, both across and within cities for the stylized model of one segment. In panels A and C, the solid downward-sloping curves describe equilibrium flows among narrow types (3). The other solid curves capture the interaction (4) between narrow and broad searchers. Panels B and D describe the flow relationship $\eta(1-I) = V$ between inventory and turnover. More narrow types (higher N) creates the dashed lines, while less instability (lower η) generates the dash-dotted lines.

higher inventory means lower turnover. Equilibrium turnover is at the point on the curve picked out by the inventory share determined in panel A.

In the data, cities with lower inventory have more searchers and less turnover. In the absence of broad types, these patterns require the comovement of two forces. First, some cities must be more "popular," in the sense that more (narrow) agents N are interested in living there. This force is essential to generate the observed variation in the number of searchers per house, $\sigma = (N-H)/H$. Higher N also affects the equilibrium condition (3): if there are more narrow types, it becomes harder to find a house at a given inventory, so the number of narrow searchers increases. Both curves in panel A of Figure 4 thus shift to the right and are now dashed; the equilibrium of a city with higher N is marked by stars, with lower inventory and more search activity.

Variation in popularity, measured by N, is thus consistent with cities lining up along a downward-sloping Beveridge curve. By itself, however, it is not consistent with the positive comovement of inventory and turnover that we see in the data. Indeed, higher N does not affect the relationship between inventory and turnover

$\eta(1-I) = V$ in panel B. Reading across from the new star equilibrium in the left panel to the right we obtain that more popular cities see lower inventory and more turnover. Intuitively, a larger pool of agents looking for houses not only reduces inventory but also allows for faster matching.

To account for the comovement of inventory and turnover in the data, we need a second force: more popular cities must also be "more stable," which means that houses come on the market at a slower rate η. The dash-dotted lines show the effect of a lower η, and the new equilibrium is marked by squares. If agents become unhappy at a slower rate, then for a given pool of searchers we have less inventory (the square in panel A) and lower turnover (the square in panel B). The higher stability in more popular cities thus contributes to the downward-sloping Beveridge curve, while helping to explain the positive comovement of inventory and turnover.

Broad Searchers and Competition within Cities.—To study the cross section of segments in a city, we now allow for "broad types" who are also interested in other segments. However, at any instant, broad types can try to buy in only one segment. The number of broad types who want to buy in the segment under study is determined by a nonnegative and increasing function of inventory, $B^B(L)$. This function captures the sensitivity of broad searchers to local conditions: a segment with higher inventory attracts more broad types wanting to buy there, consistent with the evidence in Section IIB. The idea is that, all else equal, the more houses are available for sale in a segment, the more likely it is that a broad searcher will find her favorite house in that segment.[8]

How can we characterize the equilibrium with broad searchers? The key new feature is that the number of narrow searchers B^N is no longer exogenously given: it is instead determined by the equilibrium interaction between broad and narrow types. As before, the steady-state distribution of agents across individual states must be constant, but it now includes the number of broad types in each state. Equilibrium flows must thus be consistent with broad types trading only with broad types, and narrow types only trading with narrow types, as described by condition (3), which thus continues to hold. Moreover, the ratio of total buyers B to owners H must be the same as the buyer-owner ratio for narrow types only:

$$\frac{B}{H} = \frac{B^N + B^B(L)}{H} = \frac{B^N}{N - B^N}. \tag{4}$$

We thus obtain a positive relationship between L (and thus I) and B^N that reflects *competition between broad and narrow types*. Higher inventory L attracts more broad searchers who crowd out ownership from narrow searchers. This increases

[8] In the simple model here, we assume that $B^B(L)$ does not depend directly on η, but only responds to inventory. Our quantitative model below has the same feature: it assumes that broad searchers flow to segments in their search range in proportion to inventory. In principle, it is also possible to allow $B^B(\cdot)$ to depend on η such that a segment may be less attractive to broad searchers if houses there fall out of favor more quickly. For example, a common class of models in urban economics assumes that broad searchers are indifferent in equilibrium between buying in the segment under study and receiving fixed utility elsewhere. This would imply a function $B^B(L,\eta) = q(\eta)L$, where q is the tightness of the market: the ratio of buyers to sellers. Tightness is decreasing in η: there are fewer buyers per seller if home ownership yields utility for a shorter amount of time. Online Appendix Section D shows that under plausible conditions on parameters, our results extend to models in which buyer flows also depend on η.

the equilibrium number of narrow searchers B^N. The latter continues to map directly into our measure of search activity σ, since broad searchers do not affect σ differentially across segments.[9]

Panels C and D of Figure 4 illustrate the impact of broad searchers. In panel C, equilibrium inventory and search activity are determined by the intersection of equations (3) and (4). Panel D again shows the flow condition $\eta(1-I) = V$. Section IIA documents that within most Bay Area cities, the Beveridge curve is *upward-sloping*: segments with more inventory have more search activity. In addition, those segments have more turnover. Our setup explains this pattern with the attraction of broad searchers to segments with higher inventory. The bottom row in Figure 4 holds the number of narrow searchers fixed and only varies the stability parameter η. The downward-sloping dash-dotted line in panel C is thus the same as in panel A; it shows flows among narrow types in a more stable segment (lower η).

Regardless of the clientele structure, more stability reduces inventory: in this regard, panels A and C agree. The new element in panel C is that lower inventory attracts fewer broad searchers who crowd out narrow searchers. Mechanically, we move down along the upward-sloping curve described by equation (4) to the equilibrium marked by a square. In a more stable segment, more narrow types own a house and overall search activity is lower. The more stable segment thus has both less inventory and less search activity, generating an upward-sloping Beveridge curve. Panel D in Figure 4 shows that the more stable segment also has less turnover (also marked by a square), consistent with the pattern in the data.

The Effects of Housing Policy and the Importance of Search Data.—We next illustrate the importance of search data for predicting the effects of local housing policies. Consider a researcher who tries to forecast the effects of building houses in a metro area on inventory and the pool of homeowners there. In the absence of data on housing search ranges, the typical approach is to assume that the area of interest is populated by homogeneous agents who only search in that area. With this identifying assumption, the model parameters can be inferred from observable moments such as the housing stock H, the turnover rate V, the inventory share I, and the average time it takes to find a house, which, in steady state, equals $B/(VH)$. Specifically, from our discussion of a segment with only narrow types, we have $\eta = V/(1-I)$, $N = (B/(VH))VH + H$, and $B = N - H$. It is also possible to estimate the shape of the matching function. We now ask what happens if the identifying assumption of homogeneity is incorrect because it imposes either too much or too little aggregation.

It is helpful to first calculate the effects of construction in a single segment with a fixed number of narrow searchers. Equilibrium inventory is characterized by the flow condition (3) with $B^N = N - H$. For each additional house built, the number of houses that end up in inventory is

$$\frac{dL}{dH} = \frac{L}{B}\frac{B + \varepsilon_B H(1-I)}{L + \varepsilon_L H(1-I)}, \tag{5}$$

[9] We have defined search activity σ as the (weighted) number of searchers that include a segment in their search range. Broad searchers who include many segments in their search range therefore have the same direct impact on σ across all segments in their range. Any differences in σ across segments are thus due to differences in the number of narrow searchers, B^N.

where ε_B and ε_L correspond to the elasticities of the matching function $m(B,L)$ with respect to buyers and listings, respectively.[10]

Inventory always increases after new construction, and more so if the matching function is relatively more responsive to the number of buyers than to the number of listings. We can simplify further if we follow the typical calibration strategy and assume a symmetric matching function as well as numbers of buyers and listings that are small relative to the housing stock. The latter property usually obtains because inventory shares are small in the data, and average buyer search time $B/(VH)$ is similar in magnitude to average seller time on market, I/V. With these simplifications, both elasticities are equal to one-half and the second fraction is close to 1.

We thus obtain a very simple approximate formula for the impact of new construction: for each new house built, L/B houses end up in inventory. If inventory in the initial equilibrium is exactly equal to the number of buyers, then new construction does not lead to more occupied houses: it simply leads to more inventory. If $L > B$, inventory increases by more than the housing stock: adding too many houses "clogs" the market and makes it more difficult for sellers to find buyers. Only if $L < B$ does not all of the new construction end up in inventory. So far in this section, we have kept the model at a reduced-form level without introducing preferences and welfare. However, under additional assumptions, we can use the number of houses not on the market $H - L$ as an index of consumer welfare. Indeed, suppose that agents have quasilinear preferences over housing and a numéraire good, and that they receive utility from their house only before they receive a moving shock, as in our quantitative model below. Pareto optima then maximize the sum of utilities across agents less construction costs. The sum of utilities is proportional to $H - L$, so construction improves welfare only if $L < B$.[11]

Consider now a researcher who assumes too much aggregation. For a stark thought experiment, suppose that a metro area is actually a collection of many unconnected segments, but the researcher incorrectly assumes that it is perfectly integrated, that is, he assumes that all segments are jointly searched by one broad type. Using this identifying assumption, the researcher can estimate the aggregate listings-buyers ratio L/B for the entire metro area. He can then predict the effects on inventory and welfare based on that aggregate ratio. However, imperfect integration of the area implies that the actual effect on aggregate inventory and welfare may depend on where exactly the new houses are built. For example, the researcher may measure $L/B = 1$ and conclude that there is no housing shortage in the metro area, while it may well be that some segments have $L/B < 1$, so that construction there may be welfare improving if it were sufficiently cheap.

For a second thought experiment, consider the opposite situation: an area is actually fully integrated, and a researcher focuses on one segment assuming that

[10] The flow equation $m(N-H,L) = \eta(H-L)$ determines the number of houses in inventory L as a function of the housing stock H. By applying the implicit function theorem, we get

$$\frac{dL}{dH} = \frac{\eta + m_1}{\eta + m_2} = \frac{\eta + \frac{m_1(N-H)}{m}\frac{m}{N-H}}{\eta + \frac{m_2 L}{m}\frac{m}{L}} = \frac{1 + \varepsilon_B \frac{H-L}{B}}{1 + \varepsilon_L \frac{H-L}{L}} = \frac{L}{B}\frac{B + \varepsilon_B(H-L)}{L + \varepsilon_L(H-L)},$$

where the elasticities are $\varepsilon_B = m_1(N-H)/m$ and $\varepsilon_L = m_2 L/m$.

[11] If welfare $W = H - L$, then $dW/dH = 1 - dL/dH$, which, according to equation (5), is only positive if $dL/dH \approx L/B < 1$.

it is not connected to any other segment. Using this identifying assumption, the researcher estimates a segment-specific ratio L/B and uses this ratio to infer a number of narrow searchers N. However, we are actually in a version of the model above with $N = 0$ and only broad searchers who flow to segments according to $B^B(L)$. As a result, if the segment studied by the researcher is small relative to the rest of the city, we would expect the inventory share and turnover rate to be entirely unaffected by the construction of new homes. The new houses are bought by broad searchers without affecting local conditions. Formally, with only broad searchers, inventory is determined by $\eta(H-L) = m\big(B^B(L), L\big)$. Since the segment under study is small, changes in its scale do not affect the inventory share, but only the number of broad searchers who enter.

This thought experiment illustrates the pitfalls of focusing on inventory to guide housing policy. It is tempting to view low inventory in some segment as a "housing shortage" that makes building there particularly important.[12] With narrow searchers and a downward-sloping Beveridge curve, low inventory indeed flags a large number of agents who desire local housing. With broad searchers and an upward-sloping Beveridge curve, however, low inventory may simply reflect that few houses come on the market. There is no obvious reason to build in low-inventory segments since broad searchers have no particular preference to live there.

Direct data on the pool of searchers help to distinguish the two cases and can hence improve policy decisions. Indeed, a sample of email alerts would reveal the relative share of narrow searchers B^N/B. Based on local conditions, given by equation (3), this additional information identifies N and hence puts restrictions on the function $B^B(L)$.[13] In the absence of such data, the quantitative results below suggest that a model with only narrow (broad) searchers will provide a reasonable approximation of the searcher pool across (within) cities.

IV. A Quantitative Model with Multiple Segments

In this section, we quantify the effects discussed above. To this end, we specify a fully fledged housing search model with multiple segments. At the segment level, equilibrium actions in this model imply the same flow equations as in the simple single-segment model from Section III for particular buyer flow functions $B^B(L)$ that accommodate the rich clientele structure in our data. The model here also makes explicit how transactions occur and how prices are determined. Online Appendix Section D shows the robustness of the inference to alternative modeling assumptions of the search behavior, the flow of broad buyers, and the price-setting mechanism.

Segments and Preferences.—We continue to use the notation we introduced to present the facts in Section IIA, and which we summarized in Table 1. There are $\mathcal{H}$ market segments with mass $H(h)$ houses in segment h; the mass of houses in

[12] For example, Lawrence Yun, "Four More Years? No Quick End in Sight for US Housing Shortage," *Forbes,* December 13, 2016; and Paul Davidson, "Housing Shortage Eases in Some Markets," *USA Today,* September 26, 2016.

[13] For example, if $B^B(\cdot)$ is the same for all segments, then it can be traced out from the cross section of I, V, V/B, and B^N/B. In fact, in this case the supply and demand parameters η, $B^B(\cdot)$, and N can be identified without taking a stand on the shape of the matching function.

the entire Bay Area is normalized to 1. Agent type θ is identified by search range $\tilde{\mathcal{H}}(\theta) \subset \mathcal{H}$. Search ranges are part of preferences, and type-θ agents never enjoy a house outside of $\tilde{\mathcal{H}}(\theta)$. We use the measure μ on the set of types Θ to count the number of agents of each type. The total number of agents is $\bar{\mu} = \sum_{\theta \in \Theta} \mu(\theta) > 1$. The clientele $\tilde{\Theta}(h)$ of segment h is the set of all agents who are interested in segment h, as defined in equation (1). The inventory and housing stock scanned by type θ are $\tilde{L}(\theta)$ and $\tilde{H}(\theta)$, respectively.

Agents live forever, discount the future at rate r, and receive quasilinear utility from a numéraire good and housing services. Agents only obtain housing services from a "favorite" house. After an agent moves into his favorite house in segment h, he obtains housing services $v(h) > 0$. Houses fall out of favor at rate $\eta(h)$ and then no longer provide housing services. Agents can put a house up for sale at no cost. Once the house is sold, agents search for a new house, again at no cost. Sellers incur a proportional cost c upon sale of a house. Matching in segment h occurs at the rate $\tilde{m}(B(h), L(h), h)$, where $\tilde{m}$ exhibits constant returns to scale in its first two arguments.[14]

How do agents decide on a favorite house within their search range $\tilde{\mathcal{H}}(\theta)$? Our approach is guided by the evidence in Section IIB: interest in individual segments within a search range is proportional to segment inventory. We thus assume that agents are equally likely to "fall in love" with any house for sale in $\tilde{\mathcal{H}}(\theta)$. The number of buyers $B(h)$ in segment h is

$$B(h) = \sum_{\theta \in \tilde{\Theta}(h)} \tilde{B}(\theta) \frac{L(h)}{\tilde{L}(\theta)}, \tag{6}$$

where $\tilde{B}(\theta)$ is the number of type-θ buyers. For a narrow type θ who considers only segment h, we have $L(h) = \tilde{L}(\theta)$, so this buyer contributes one-for-one to $B(h)$. To determine the contribution to $B(h)$ of a broad type θ, we multiply the number of these buyers by $L(h)/\tilde{L}(\theta)$, which represents the share of total inventory scanned by type θ that is in segment h.

Expression (6) is the counterpart to equation (4) in the one-segment reduced-form model from Section III. In that model, the spirit of our buyer-flow assumption would be captured by the functional form $B^B(L) = qL$ for some exogenous constant q. In the current model, segments differ in their clientele structure. As a result, there are many segment-specific measures of the sensitivity of a segment's buyer pool to segment inventory, $q(h) = \sum_{\theta \in \tilde{\Theta}(h)} \tilde{B}(\theta)/\tilde{L}(\theta)$, that are all jointly determined in equilibrium.

Matching, Bargaining, and Equilibrium.—Once a buyer and seller have matched, the seller makes a take-it-or-leave-it offer. If the buyer rejects the offer, the seller keeps the house and the buyer continues searching. If the buyer accepts the offer, she pays the offer price. The seller receives the offer price net of the proportional cost c which goes to a real estate agent. The seller then starts to search, whereas the buyer moves into the house and begins to receive utility $v(h)$.

[14]The matching function is increasing in the number of buyers and sellers and satisfies $\tilde{m}(0, L, h) = \tilde{m}(B, 0, h) = 0$. It is also allowed to depend on the segment h directly (other than through the number of buyers and inventory). For example, the process of scanning inventory could be faster in some segments because the properties there are more standardized, or because more open houses are available to view properties.

An equilibrium is a collection of agent choices such that each agent chooses optimally given the distribution of others' choices. In particular, owners decide whether to put their houses on the market, sellers choose price offers, and buyers choose whether to accept those offers. We focus on steady-state equilibria in which (i) owners put their house on the market if and only if their house has fallen out of favor, so that the owners no longer receive housing services from it, and (ii) all offers are accepted.

The model endogenously determines inventory shares $I(h)$ and turnover rates $V(h)$ for each segment. It also determines the number of searchers $\tilde{B}(\theta)$ of each type, which allows us to calculate segment-level search activity $\sigma(h)$ as defined by equation (2). The cross section of these observables is shaped by three distinct forces. Supply is represented by the rate $\eta(h)$ at which houses fall out of favor. Demand is captured by the clientele structure, the distribution of types $\mu(\theta)$. Finally, the segment-specific effect on match rates $\tilde{m}(\cdot,\cdot,h)$ represents differences in market frictions across segments. The next section describes how data on the observable endogenous variables $I(h)$, $V(h)$, and $\tilde{B}(\theta)$ allow us to quantify these three forces.

A. *Housing Demand and the Cross Section of Housing Markets*

Our quantitative analysis proceeds in two steps. First, we show how the cross section of inventory, turnover, and search activity in the Bay Area reflects housing demand (in particular, the presence of broad searchers) as well as the other two exogenous forces capturing supply and market frictions. We provide additional evidence that validates and helps interpret the demand estimates. Section V then studies price formation and conducts counterfactuals.

Identification.—The analysis in this section requires only the supply and demand parameters; it does not depend on the details of bargaining or the shape of the matching function. The intuition is as in Section III. With a fixed number of houses and agents, the steady-state distribution of agent states (searching for a house, listing one for sale, and owning without listing) is pinned down by house and agent flows, regardless of pricing. Moreover, matching frictions matter for agent flows only via the rates at which buyers find houses in a given segment, defined as $\alpha(h) = m(h)/B(h)$.

More formally, our identification result in online Appendix Section F establishes a one-to-one mapping between two sets of numbers. The first set consists of the supply and demand parameters, $\eta(h)$ and $\mu(\theta)$, as well as the vector of house finding rates $\alpha(h)$. The second set consists of objects we observe in the data: the inventory share $I(h)$, the turnover rate $V(h)$, the relative frequencies of search ranges $\tilde{B}(\theta)$, and the *average* time it takes for a buyer to find a house.[15] We use this mapping to back out the vectors of $\eta(h)$, $\mu(\theta)$, and $\alpha(h)$.

[15] While the equilibrium depends on the total number of buyers of each type, $\tilde{B}(\theta)$, we only observe them up to a constant: the email alert data allow us to infer the relative number of each type, $\tilde{B}(\theta)/\bar{B}$, but we have no information on the overall number of buyers, $\bar{B}$. As an additional target moment, we thus set the average match rate for buyers to 20 percent per month, corresponding to the average match rate for inventory in our data. The average time it takes for a buyer to find a house is therefore about five months. This choice does not affect the relative behavior of market and search activity across segments, and is thus not important for most of our results.

We refer to the supply parameter $\eta(h)$ as the *instability* of the segment: in a more unstable segment, houses come on the market at a faster rate. To measure the *popularity* of a segment, we define the weighted number of agents who are interested in a house there:

$$\pi(h) := \frac{1}{H(h)} \sum_{\theta \in \tilde{\Theta}(h)} \mu(\theta) \frac{H(h)}{\tilde{H}(\theta)} = \sum_{\theta \in \tilde{\Theta}(h)} \frac{\mu(\theta)}{\tilde{H}(\theta)}. \tag{7}$$

Note that this definition weights types of agents by their share in the overall population (which is not directly observable), while search activity $\sigma(h)$ in equation (2) weights them by their share among searchers (which we could observe directly in the email alert data). Popularity is thus an exogenous determinant of demand for segment h; it only depends on the distribution of types $\mu(\theta)$. In contrast, $\sigma(h)$ is determined endogenously and depends on the equilibrium share of agents of each type that are searching at any point in time. We continue to count individuals who are potentially interested in many segments less toward the popularity of segment h than individuals who only like that particular segment.

Parameter Estimates.—Table 4 summarizes our estimates of instability $\eta(h)$, popularity $\pi(h)$, and the house finding rate $\alpha(h)$. Panel A provides information on the distribution of the estimates across segments. Panel B reports correlations both among the estimates themselves as well as between estimates and observables. Here we compare variables across all segments, across all cities, and across segments within the city of San Francisco.

Instability $\eta(h)$, which captures the frequency of moving shocks in a segment, tracks turnover almost exactly, and its moments in Table 4 are essentially the same as those reported for the turnover rate in Table 2. The result follows from the flow condition $\eta(h)(1 - I(h)) = V(h)$ for each segment, together with the summary statistics in Table 2. Inventory shares are so small, their seventy-fifth percentile is at 1.51 percent, that $\eta(h) \approx V(h)$. Intuitively, because a house remains on the market for a much shorter time than it is occupied by an owner, turnover is almost entirely accounted for by the frequency of moving shocks.

Popularity at the segment level, $\pi(h)$, ranges between 0.20 and 2.39, with an interquartile range of 0.79 to 1.14. The fact that popularity is below one for many segments is indicative of the importance of partial integration. If segments were either perfectly segmented or perfectly integrated, the number of (weighted) people interested in each segment would be larger than the number of houses, and popularity would have to be above 1. However, segments that are only considered by relatively broad searchers can have a popularity less than 1, because nobody who lives there has a particular preference for that location. Popularity also comoves strongly with search activity at all levels of aggregation: in equilibrium, more popular segments generally have more (weighted) searchers per house.

The Cross Section of Cities.—The middle rows of panel B in Table 4 show how the forces from Figure 4 quantitatively account for market activity in the cross section of cities. Instability generates comovement of inventory and turnover; in fact, it is almost perfectly correlated with both. At the same time, more popular cities see

Table 4—Quantitative Results from Estimating the Model

	Estimated parameters			Hypothetical perfectly segmented benchmark			
	$100 \times \eta(h)$	$\pi(h)$	$\alpha(h)$	$\hat{\sigma}(h) - \sigma(h)$	$\hat{\pi}(h) - \pi(h)$		
Panel A. Estimated parameters versus hypothetical segmentation							
Mean	0.24	0.99	0.37	0.02	−0.02		
Q25	0.16	0.79	0.14	−0.17	−0.21		
Q50	0.21	0.98	0.24	−0.05	−0.03		
Q75	0.29	1.14	0.44	0.12	0.13		
	$\eta(h)$	$\pi(h)$	$\alpha(h)$	$I(h)$	$V(h)$	$\sigma(h)$	$\log(p(h))$
Panel B. Correlation between estimated parameters and data moments							
Across segments							
$\eta(h)$	1	−0.01	0.23	0.93	1.00	0.01	−0.48
$\pi(h)$		1	−0.53	−0.04	0.00	0.71	0.07
$\alpha(h)$			1	0.27	0.23	−0.43	−0.22
$\hat{\sigma}(h) - \sigma(h)$	0.44	0.09	−0.04	0.45	0.44	0.28	−0.16
Across cities							
$\eta(h)$	1	−0.23	0.52	0.93	1.00	−0.24	−0.55
$\pi(h)$		1	−0.60	−0.27	−0.23	0.66	0.35
$\alpha(h)$			1	0.59	0.52	−0.70	−0.65
$\hat{\sigma}(h) - \sigma(h)$	0.37	−0.02	0.00	0.40	0.37	0.09	−0.11
Within San Francisco							
$\eta(h)$	1	0.29	−0.23	0.87	1.00	0.48	−0.24
$\pi(h)$		1	−0.60	0.34	0.29	0.90	−0.01
$\alpha(h)$			1	−0.46	−0.22	−0.60	0.22
$\hat{\sigma}(h) - \sigma(h)$	0.83	0.18	−0.15	0.81	0.83	0.33	−0.12

Notes: The estimated parameters from the model exactly match three moments from the data: the inventory share $I(h)$, the turnover rate $V(h)$, and the relative frequency of search ranges $\beta(\theta)$. The total number of buyers $\bar{\mu} - 1$ is set to match a five-month average buyer search time.

more searchers per house and lower inventory. Differences in popularity are thus a key force behind the downward-sloping Beveridge curve across cities. The positive impact of popularity on turnover is muted by the fact that more popular cities are also more stable.

In addition to the supply and demand effects that were already present in the simple model in Section III, the quantitative model also identifies significant differences in matching frictions across cities. Indeed, we find that in cities with low inventory and turnover but many searchers, matching is particularly slow (that is, $\alpha(h)$ is low).[16] Intuitively, cities with many searchers must be popular. Slow matching explains why this popularity does not translate into high turnover.

The Quantitative Importance of Partial Integration on Estimates of Segment Popularity.—To develop a summary statistic for the quantitative effect of broad searchers, we compare the actual economy, which exhibits partial integration, with a hypothetical perfectly segmented benchmark economy. We construct the latter by changing demand parameters so as to remove integration, holding all other

[16] To illustrate how the matching technology affects $\alpha(h)$, consider a Cobb-Douglas matching function $m(h) = \bar{m}(h)B(h)^{\delta}L(h)^{1-\delta}$, which implies $\log\bar{m}(h) = \delta\log\alpha(h) + (1-\delta)\log(V(h)/I(h))$.

parameters fixed. At the same time, we require that the benchmark economy still matches the same observed inventory shares and turnover rates as the actual economy. This pins down a unique vector of hypothetical demand parameters $\hat{\mu}$.

The benchmark economy can be viewed as the model estimated by an econometrician who observes the inventory share $I(h)$, the turnover rate $V(h)$, and buyer match rate $\alpha(h)$ by segment, but who does not have information on integration and proceeds by assuming that the economy is perfectly segmented. This misspecification will lead the econometrician to infer the wrong demand parameters, and hence incorrect measures of popularity $\hat{\pi}(h)$. We thus use the distribution of the difference $\hat{\pi}(h) - \pi(h)$ to assess the size of the specification error from ignoring partial integration.[17]

Table 4 shows that the error incurred by an econometrician who ignores partial integration is large. Quantiles for the difference in popularities $\hat{\pi}(h) - \pi(h)$ across segments are reported in the rightmost column in panel A. The interquartile range of this difference is of the same order of magnitude as the interquartile range of the estimated parameter $\pi(h)$ itself. The reason is that the range of estimated popularities is much narrower for the hypothetical fully-segmented benchmark economy: the twenty-fifth percentile for $\hat{\pi}(h)$ is at 1.005, while the seventy-fifth percentile is at 1.015.

Intuitively, the econometrician infers too little dispersion in popularity across segments because he ignores the endogenous response of broad searchers to market conditions. A model with narrow searchers must explain observed market and search activity entirely via narrow searchers' demand, as opposed to broad searchers chasing inventory. In the actual economy, some unpopular and unstable segments attract very few narrow searchers, but nevertheless see substantial search activity due to attention from broad searchers who are drawn to the segment's high inventory. Ignoring partial integration therefore overstates the popularity of such unstable high-inventory segments. As discussed in Section III, overestimating the popularity of a segment generates misleading predictions on the quantity-effects of additional construction, changes in zoning regulation, public transit investments, and other place-based policies. We discuss below how it will also bias estimates of the effects of such policies on equilibrium transaction prices.

Broad Searchers and the Beveridge Curve within Cities.—We next quantify the contribution of broad searchers to the shape of the Beveridge curve within cities. In particular, the perfectly segmented benchmark economy in the previous section predicts a different cross section of search activity, $\hat{\sigma}(h)$, and hence a different Beveridge curve. We can therefore use the difference in search activity $\hat{\sigma}(h) - \sigma(h)$ to summarize the effect of partial integration on the Beveridge curve. If integration was not important, then the actual and hypothetical Beveridge curves would coincide.

The key difference between the benchmark and actual Beveridge curves is that the slope of the latter partially reflects the response of broad searchers to inventory.

[17] The other parameter estimates are not materially affected by misspecification. We have seen that the parameter $\eta(h)$ closely tracks turnover: this result follows from the flow conditions regardless of the clientele patterns. Moreover, given the equilibrium conditions and the fact that the matching function remains unchanged, the benchmark economy predicts the same number of buyers and buyer match rates $\alpha(h)$ as the actual economy.

In high inventory segments, broad searchers crowd out narrow searchers which increases the overall number of searchers per house. As a result, the actual Beveridge curve is steeper than the benchmark curve in the (σ, I)-plane and the difference $\hat{\sigma}(h) - \sigma(h)$ should be positively correlated with inventory in integrated areas.

Table 4 shows that the contribution of broad searchers to the slope of the Beveridge curve within San Francisco is large. The correlation coefficient of 0.81 between $I(h)$ and $\hat{\sigma}(h) - \sigma(h)$ implies that a 1 standard deviation increase in inventory increases $\hat{\sigma}(h) - \sigma(h)$ by 0.54 standard deviations of the search activity measure $\sigma(h)$ itself. Within San Francisco, a perfectly segmented benchmark that ignores broad searchers thus generates a Beveridge curve that is much flatter then the true Beveridge curve we measure. Across cities, in contrast, broad searchers matter less: ignoring the presence of broad searchers adds only about 10 percent worth of search activity per unit of inventory. This finding is consistent with the results in Section IIA that showed less integration across cities.

What exogenous forces generate the differences in inventory that attract the broad searchers? In principle, differences in instability, popularity, or matching frictions could all generate differences in inventory across segments. Quantitatively, however, the flow of broad searchers is directed to a large extent by differences in instability. Indeed, the correlation of $\hat{\sigma}(h) - \sigma(h)$ with $\eta(h)$ is much larger than that with $\pi(h)$ or $\alpha(h)$. We can therefore sum up the key mechanism as follows: in more unstable segments, more houses come on the market. If these segments are part of an integrated area, then broad searchers are attracted to the higher inventory and crowd out narrow searchers, generating an upward-sloping Beveridge curve.

San Francisco Bay Area versus Other Metro Areas.—The key intuition described above is that the endogenous flow of broad searchers to high-inventory segments within their search ranges can induce an upward-sloping Beveridge curve across market segments with a large common clientele. Our search data reveal that most housing search in the Bay Area is along city lines: 61 percent of search queries specify a city as the finest geographic unit. This implies a large common search clientele within Bay Area cities, and rationalizes the upward-sloping Beveridge curve across segments within most cities. However, in other parts of the United States, the geographic and political units that are jointly searched can potentially differ. For example, in Massachusetts, "cities" are much smaller political units, and different cities such as Cambridge and Somerville are probably regularly searched jointly. This generates the potential for an upward-sloping Beveridge curve across segments in those units. On the other hand, in New York City, searchers are unlikely to consider the entire city, and might only search parts of different boroughs jointly. This suggests we are less likely to see an upward-sloping Beveridge curve across all New York City submarkets than we are to see it across submarkets within the same borough.

B. *Further Evidence on Housing Demand*

Our estimation infers housing demand from data on market and search activity. To help with interpreting these demand estimates, we now relate our estimates of popularity, our segment-level summary statistic for demand, to observable characteristics of segments. We also relate search breadth, our key individual-level statistic, to searcher demographics. Finally, we compare model-implied household flows between segments to data on actual moves, providing an overidentifying restrictions test on the structure of our model.

Popularity and Segment-Level Observables.—What makes a segment popular? And how can policymakers and researchers identify popular segments in the absence of detailed search data? One characteristic associated with popular cities appears to be the service flow from properties in the segment, which is a proxy for their quality. In particular, Table 4 shows a 35 percent correlation between log price and popularity across cities. More expensive cities thus have a larger average clientele interested in living there.

When comparing segments within cities, the correlation between price and popularity is much weaker. To understand which other characteristics are correlated with segment popularity, we construct the housing-stock-weighted average popularity of all segments in a zip code. We then use a simple regression to relate this zip-code-level popularity measure to characteristics observable at the zip code level: school quality, the availability of restaurants and bars, crime levels, and weather conditions.[18]

The results in Table 5 show that the availability of bars and restaurants is the most important correlate of popularity at the zip code level: a 1 standard deviation increase in the number of restaurants and bars is associated with an increase in popularity of 0.145, or 0.54 standard deviations, and the number of bars and restaurants explains 24.7 percent of the across-zip-code heterogeneity in popularity. The strong correlation between the number of restaurants and popularity is true both unconditionally, as well as when comparing zip codes within cities. Other zip code characteristics also matter, but they are quantitatively less important. A 1 standard deviation increase in school quality increases popularity by 0.04, whereas a 1 standard deviation increase in violent crime reduces popularity by 0.04. The weather in a zip code is also correlated with popularity: more rain, more extreme hot weather, and more extreme cold weather all reduce popularity, with more extreme hot weather having the largest effect. These correlations suggest that using zip-code-level characteristics such as the number of restaurants and bars might allow policymakers who do not observe housing search data to better target localized housing policies (such as additional construction) to particularly popular segments.

Search Breadth and Demographics.—Who are the broad searchers? While we do not directly observe demographics for the individuals setting email alerts in our data, our model estimates imply a distribution of search breadths for households living in equilibrium in each segment. We can thus compare the average search breadth

[18] We measure school quality as the student-weighted average Academic Performance Index (API) across all schools in a zip code, as reported by the California Department of Education. To measure the availability of restaurants, we divide the number of establishments with SIC code 58 (eating and drinking places) by the number of housing units. Crime levels are measured on a scale of 0–100, as provided by Bestplaces.net. To measure weather, we calculate the total number of inches of rain, the total number of cooling degree days, and the total number of heating degree days, as reported by Melissa Data. Heating degree days (HDD) and Cooling degree days (CDD) are measures of how far, and for how long, temperatures deviate from 70 degrees Fahrenheit. For example, every day the temperature is at 65 degrees Fahrenheit counts as 5 HDD.

Table 5—Drivers of Zip-Code-Level Popularity

	Coefficient	Effect of 1 SD increase on zip code popularity	R^2 in univariate regression (%)
Restaurants and bars (per 100 housing units)	0.145 (0.027)	0.14	24.7
School quality (average API /100)	0.050 (0.026)	0.04	2.6
Violent crime (scale 0–100)	−0.003 (0.001)	−0.04	0.5
Rain (inches)	−0.001 (0.000)	−0.02	0.0
Cooling degree days (k days)	−0.109 (0.050)	−0.03	2.3
Heating degree days (k days)	−0.052 (0.029)	−0.03	0.2

Notes: The first column shows the coefficient estimates from a multivariate ordinary least squares regression of zip-code-level popularity on observable zip code characteristics. Observations = 183. Robust standard errors in parentheses. The multivariate regression has an R^2 of 32.2 percent. The second column shows the implied effects of a one standard deviation increase in the characteristics on zip code popularity. The across-segment standard deviation of popularity is 0.26. The final column shows the R^2 from univariate regressions of popularity on each of the zip-code-level characteristics.

of people living in a zip code, measured by the share of Bay Area inventory scanned, with demographic information from the five-year estimates of the 2012 American Community Survey.[19]

We focus on residents' age, income, and the presence of children. Figure 5 presents binscatter plots to show the relationship between these demographic measures and search breadth. In panel A, we group our 191 zip codes by the median age of the inhabitants. There is a strong negative relationship between age and search breadth. People living in zip codes with a median age of 30 search five times as much inventory, on average, as people living in zip codes with a median age of 50, and differences in median age explain almost one-third of the across-zip code heterogeneity in the average search breadth of residents. Panel B shows that the average search breadth of households living in zip codes with many children is higher, and panel C shows that people living in zip codes with higher median incomes have smaller search ranges. All these relationships are true both unconditionally and conditional on the other two demographic measures.

While our paper takes individuals' search ranges as the primitive representation of their preferences rather than deriving them as part of an optimization problem, there are a number of ways to rationalize the observed correlations between search breadth and searcher characteristics. For instance, to the extent that older and richer people perceive a higher marginal value of time (for example, because they earn a higher wage), this would reduce the distance that these agents are willing to commute to work, and naturally lead to a narrower optimal search range. We view it as a promising area for future work to endogenize buyer search ranges.

[19] The average search breadth of people *living* in a zip code in equilibrium can be different from that of people *searching* in a zip code, which we explored in Figure 3.

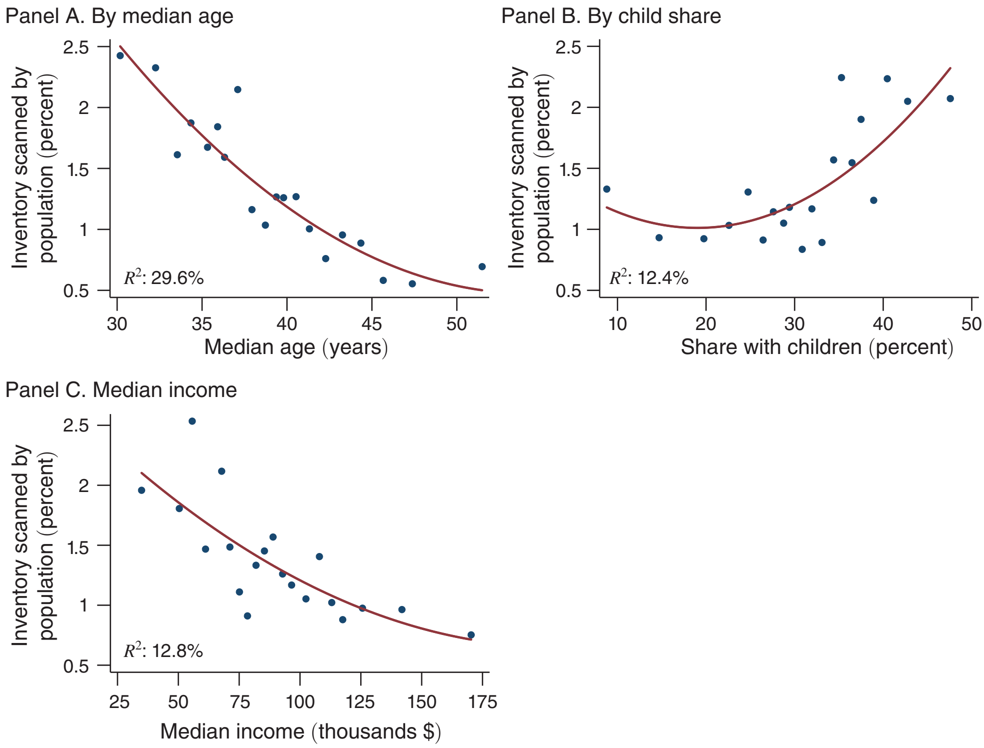

Figure 5. Search Breadth and Demographics

Notes: Figure shows binscatter plots of the average search breadth of people living in a zip code in equilibrium (measured as the average share of Bay Area inventory scanned), separately by zip code demographics. We show search breadth by median age (panel A), by the share of population with children (panel B), and by median zip code income (panel C).

Medium-Run Endogeneity of Search Ranges and Segment Characteristics.—The focus of our paper is to explore the steady-state relationships between housing search patterns and segment-level equilibrium housing market outcomes. This allows us to study how, at any point in time, the distribution of housing market activity across segments is shaped by the interaction of heterogeneous search clienteles, and how these housing market outcomes might respond to small shocks to local supply or demand conditions. Our setup is not designed to study a potential endogenous medium-run feedback between changes in segment-level searcher clientele, segment characteristics, and housing market activity, that might arise, for example, due to the gentrification of neighborhoods. We think that the study of such medium-run transitions between different steady states is an interesting area for further research.

Moving Patterns in Model versus Data.—While our estimation targets only moments of the cross section of housing market activity, the model also has implications for the flow of households between segments. We now confront those implications with data on actual household moves. Such an evaluation of model predictions that were not targeted in the estimation provides a joint test of our model assumptions and the quality of our search data. To measure the flow of households between segments, we use a sample of all individuals who moved to a new Bay Area address between May 2012 and October 2012. The data come from Acxiom, a marketing analytics company that compiles this list of movers from Change of Address notices filed with the US Postal Service.

Our sample contains 96,170 individuals moving to a new address in one of the 191 zip codes in our sample; for these individuals, we have information on their new and previous addresses. We focus on movers who had previously also lived in one of the 191 Bay Area zip codes, about 75 percent of the sample. We then compute the shares of movers between each of the 191×191 (directed) zip code pairs in the data and compare them to the shares of movers predicted by the model, which we obtain by aggregating model-implied segment-to-segment flows to the zip code level.

Figure 6 shows a scatter plot of the mover share in the data and the model at the (directed) zip code-pair level. The correlation coefficient is 82 percent. The high correlation is not only driven by households moving within the same zip code: when we exclude such moves, the correlation drops only slightly, to 75 percent. We conclude that moving patterns provide additional support for our quantitative account of Bay Area housing market dynamics.

V. Prices and Spillovers

In this final section, we explore how the forces in our model relate to equilibrium prices across segments. We also explore how the clientele structure in the Bay Area shapes the responses of different segments to housing market shocks, such as the influx of new narrow searchers as a result of the gentrification of neighborhoods.

A. *Equilibrium Prices and Frictional Discounts*

Our model captures two distinct housing market frictions. The first is search: owners whose house falls out of favor spend time first looking for a buyer and then for a new house. During this time they forgo the flow utility of living in their favorite house. The second friction is the transaction cost paid upon sale. In equilibrium, both of these costs are capitalized and reduce the house price relative to a frictionless model: every buyer takes into account that both he and all potential future buyers may have to sell and hence search and pay transaction costs.

Online Appendix Section E derives a convenient approximate formula for the equilibrium price in a segment, $p(h)$, which highlights how the resulting *frictional discount* reflects both frictions:

$$p(h) \approx \frac{v(h)}{r}\big(1 - I(h)\big)\frac{r}{r + cV(h)}. \tag{8}$$

In a frictionless market, matching is instantaneous, so that there is no outstanding inventory ($I(h) = 0$), and there are no transaction costs ($c = 0$). As a result, the price simply reflects the present value of future housing services $v(h)/r$.

Search and transaction costs modify the frictionless price $v(h)/r$ by first reducing housing services proportionately by $I(h)$ and then increasing the discount rate to $r + cV(h)$. The inventory share measures the price discount due to search frictions,

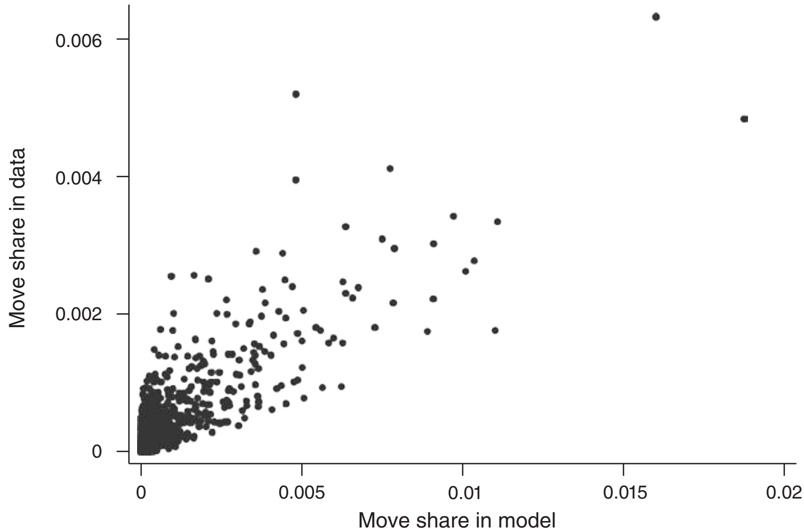

Figure 6. Moving Activity: Model versus Data

Note: Figure shows the share of total moves predicted by the model (horizontal axis) versus the moving share in the data (vertical axis) for each (directed) pair of zip codes.

and captures the fact that no household obtains housing services while a property is listed for sale. This discount is zero when matching is instantaneous. From Table 2, the size of the search discount is typically a few percentage points. The interest rate does not matter for its size (at least approximately) because time on market is fairly short. The second fraction, $r/(r + cV)$, measures the present value of transaction costs: it is zero if there is no turnover or if selling houses is costless. Here the interest rate is important: if future transaction costs are discounted at a lower rate, then the discount is larger.

Frictional Discounts by Segment.—We now ask how market frictions quantitatively affect house prices across segments. To compute the frictional discount, we need an estimate of the present value of future housing services in a segment. If we assume a real interest rate of 1 percent and a transaction cost equal to 6 percent of the resale value of the house, we can use equation (8) to back out the utility values $v(h)$ such that the model exactly matches the cross section of transaction prices, inventory shares, and turnover.

Figure 7 shows the results. Panel A plots transaction prices by segment against the frictional discounts, stated as a percentage of the frictionless price. Panel B shows the geographic distribution of the frictional discounts.

There are two notable results. First, frictional discounts are large. The median discount is 14 percent and the ninetieth percentile discount is 24 percent. From Table 2 and the approximating formula above, both search and transaction costs contribute to this result, usually in the same direction since inventory shares and turnover rates are highly correlated. However, transaction costs are quantitatively more important. While search costs generate frictional discounts of up to 6 percent, the capitalized

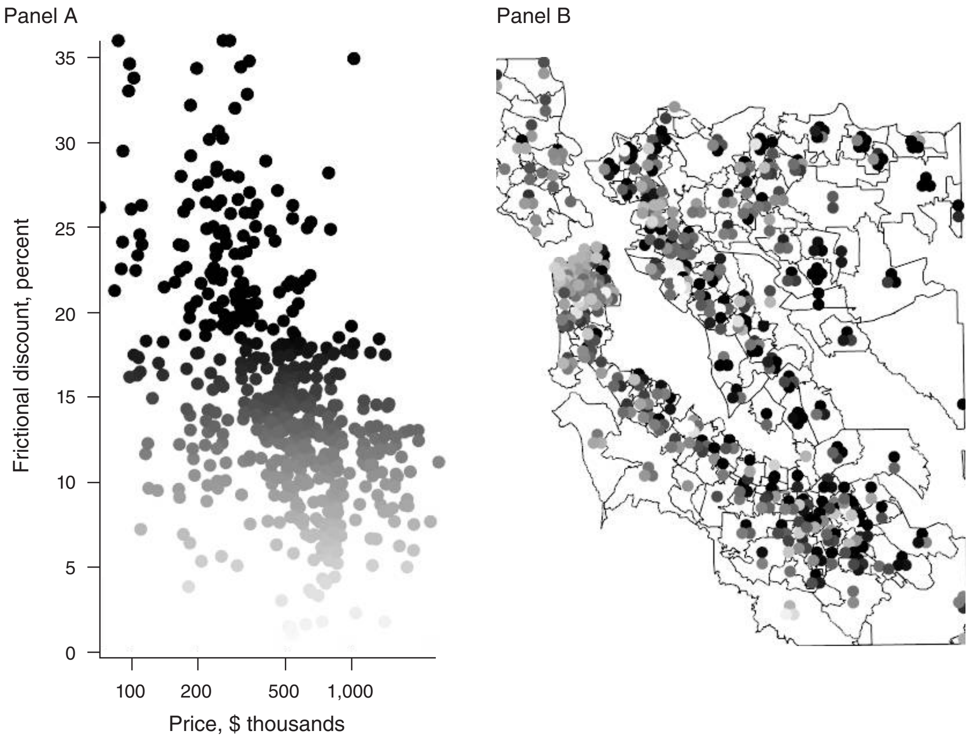

Figure 7. Frictional Discounts

Notes: Figure shows frictional discounts across segments, stated as a percentage of the frictionless price $v(h)/r$. Panel A shows the mean segment price versus the segment frictional discount in a percentage of frictionless price. Color coding reflects the size of the frictional discount. Panel B shows a map with dots for each segment in same color as in panel A. Dots for segments within the same zip code are arranged clockwise by price with the lowest-priced segment at noon.

value of transaction costs is what leads to double-digit discounts. Intuitively, the relatively small frictional discount due to search costs is because inventory shares are relatively low, and houses are therefore occupied most of the time.

The second result is that frictional discounts differ widely by segment, often within the same zip code. Table 3 shows that inventory share and turnover rate exhibit about the same amount of variation within and across zip codes. The search and transaction costs inherit these properties, respectively. In poor (low-price) segments with high turnover and high inventory, both search and transaction costs are high. As a consequence, prices are significantly lower than they would be in a frictionless market. In rich (high-price) segments with low turnover and low inventory, frictional discounts are still significant, but they are considerably smaller.

B. *Comparative Statics*

In this section, we perform comparative static exercises to show how clientele patterns matter for the transmission of localized shocks across housing market segments. Motivated by recent debates about housing shortages in parts of the San Francisco Bay Area, we ask what happens when new construction adds houses in different neighborhoods. We compare two neighborhoods that are similar in size and price, but differ in their clientele patterns. For each of these neighborhoods, we recompute the steady-state equilibrium under the assumption that 1,000 additional houses are added in that neighborhood. Since both neighborhoods consist of several segments, we allocate new houses to segments in proportion to their housing stock.

The first neighborhood is zip code 94015 in Daly City, a suburb right outside the San Francisco city limits. It contains about 11,000 houses with an average value of $480K. The average inventory is 106 houses; at our estimated parameters, this inventory is considered by 233 active searchers. We choose Daly City because narrow buyers are prevalent there. Among agents interested in Daly City, the share of narrow types, defined here as those who search in five segments or less, is 28 percent. In contrast, the share of broad types, defined as those who search in 20 segments or more, is only 15 percent. The second neighborhood is San Francisco's Outer Mission, zip code 94112. For comparability with 94015, we select only the cheapest three segments in this zip code; we thereby obtain about the same total housing stock and average price. However, the population of searchers is quite different. There are 1,054 searchers looking at an average inventory of 84 houses. Moreover, among agents interested in the Outer Mission, the share of narrow types as defined above is only 3 percent, which is tiny compared to the 74 percent share of broad types.

Figure 8 illustrates how broad searchers integrate the two neighborhoods with the rest of the San Francisco Peninsula. Panel A shows the *joint searcher share* of each segment with the Daly City zip code 94015, that is, the share of 94015 searchers who also search that segment. Darker colors indicate more integration between 94015 and the respective segment. Zip code 94015, which is the area shaded in dark gray, is most integrated with its adjacent zip code 94014. There are also weaker connections to towns to the south as well as the city of San Francisco to the north, shaded in light gray. The San Francisco segments that are searched jointly with Daly City tend to be the less expensive ones within a zip code, indicated by the top twelve o'clock dots of the circle for that zip code. Panel B of Figure 8 maps joint searcher shares with the Outer Mission. There is strong integration with cheap segments within all of San Francisco, whereas the connection to more expensive segments as well as to segments outside the city limits is much weaker.

Panels C and D of Figure 8 show the effects of new construction on inventory shares across segments. As expected from our simple model in Section III, construction increases inventory. At the same time, turnover rates remain essentially unchanged, since houses come on the market at the same rates $\eta(h)$ as before. The time to sell a house thus moves proportionally with inventory. Moreover, price effects follow from equation (8): while changes in the transaction cost discount are negligible, changes in inventory shares move the search discount. Panels C and D of Figure 8 thus also show the distribution of increases in time on market and declines in price. Darker dots indicate stronger responses in inventory shares (as well as time on market and price changes).

For both zip codes, we see the strongest responses in the zip codes where construction takes place. The main result is that a shock to a less integrated neighborhood has locally larger effects that spread less widely to other segments. In particular, the effects of construction in Daly City 94015 are disproportionately felt in 94015 itself,

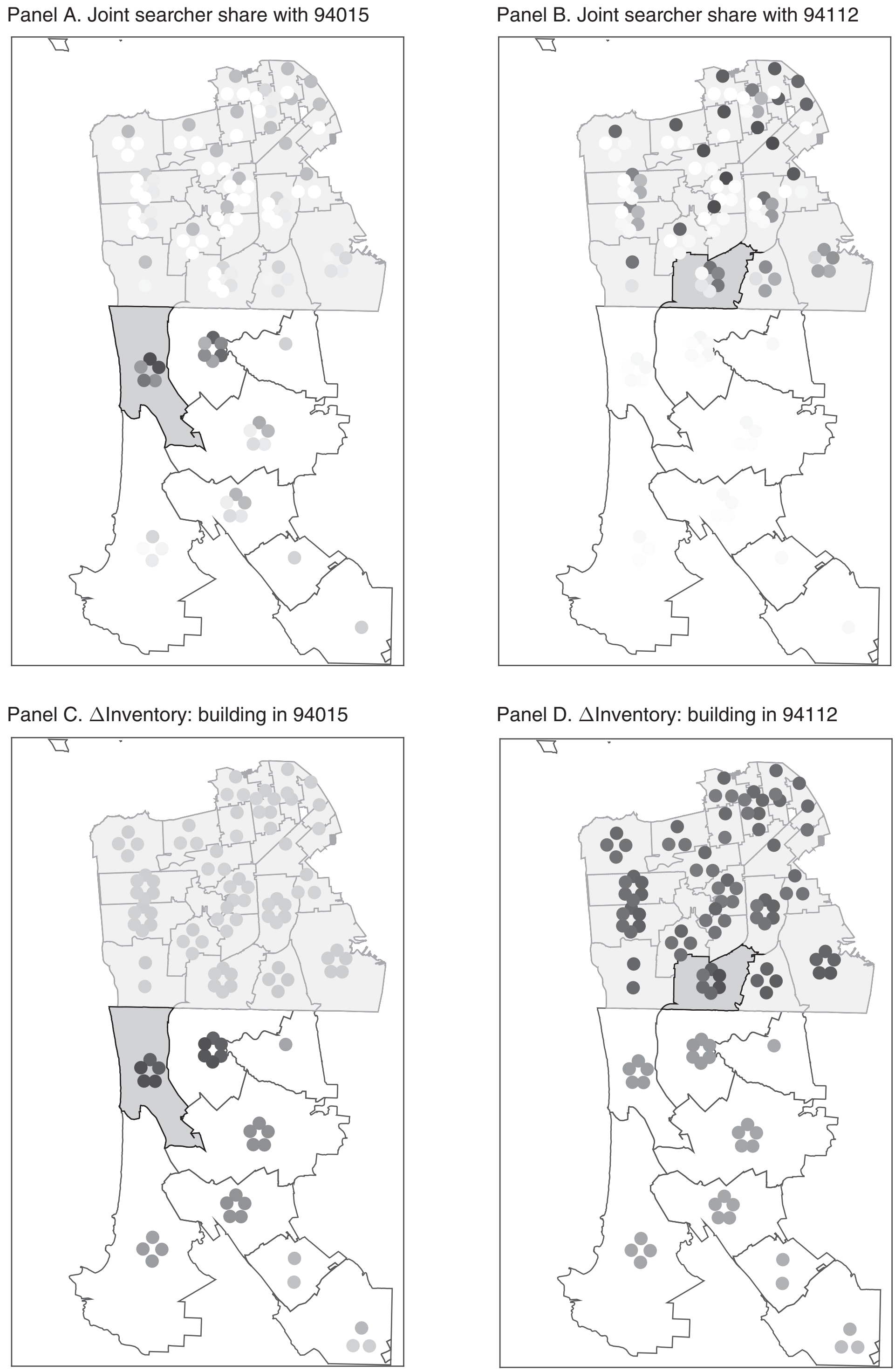

FIGURE 8. JOINT SEARCHER SHARE AND INVENTORY RESPONSE TO CONSTRUCTION

Notes: Panels A and B show the joint searcher share with 94014 and 94112, respectively. Panels C and D show the percent change in inventory share in response to the construction of 1,000 houses in zip code 94015 (Daly City) and the three cheapest segments of zip code 94112 (San Francisco Outer Mission).

as well as in the adjacent zip code 94014. Here, inventory shares increase by more than 30 percent. The effect on the typical segment in the city of San Francisco to the north, shaded in light gray, is much smaller at around 8 percent. In contrast, construction in the Outer Mission has sizable spillover effects throughout the entire city of San Francisco. In the segments where construction takes place, the average increase in inventory is only about 18 percent. However, in all San Francisco neighborhoods, the increase is 13 percent or more. Noticeably, there are similar-sized increases even in the more expensive segments in San Francisco, which do not share many direct broad searchers with the Outer Mission (see panel B). These are generated by chains of connections through other searchers who integrate cheaper and more expensive segments within a zip code. Since changes in inventory translate directly into frictional discounts, price responses are also smaller and more diluted for shocks that hit the Outer Mission, but larger and more concentrated for shocks to Daly City.

Differences in integration are also reflected in the surplus from housing that accrues to different types of agents. Indeed, construction in Daly City *lowers* the number of narrow types (defined as above) who are happy homeowners. While it has become easier to find a house, it is also harder to sell it, which may lower the overall speed at which unhappy narrow agents can again become happy owners. The effect is already present if there are only narrow searchers, as in the perfectly segmented toy example of Section IV; here it is compounded by an inflow of broad searchers from neighboring zip codes. After construction in Daly City, the number of narrow types who are "unhappy," that is, narrow types who do not receive surplus from a house, increases by 8 percent. At the same time, broad types benefit: the number of unhappy broad searchers interested in Daly City declines by 3 percent. For the Outer Mission, in contrast, the numbers of unhappy agents declines by 7 percent for both narrow and broad types interested in living there.

Information on clientele patterns is also critical for assessing the impact of construction on Bay Area aggregates. Consider a researcher who is interested only in Bay Area aggregates, and hence calibrates a model with one segment to the entire Bay Area. If our model is correct, and the researcher calibrates the number of agents to match average buyer search time, he finds the same number of buyers as we do, which is approximately the same as the number of houses in inventory. In the absence of information about clientele patterns, his predicted response to the construction of 1,000 houses is then that 1,074 more houses flow to inventory, and that welfare would decline as a result of the construction. On aggregate, a similar buyer search time and seller time on market imply that there is no shortage of housing. In contrast, our model says that building in Daly City or the Outer Mission increases aggregate Bay Area inventory by only 1,017 and 847 houses, respectively.[20] This is because those areas have local housing shortages that allow the absorption of more houses; as a result, depending on the construction costs, welfare may increase from construction in these areas.

[20] Interestingly, the local increase in inventory is quite small: within San Francisco, construction in Daly City or the Outer Mission increases inventory by 76 and 148 houses, respectively. Much of the overall increase is accounted for by small increases of one or two houses per segment across most of the Bay Area. This occurs even in the absence of broad searchers who search the entire area, via chains of connections by more narrow searchers.

The experiments also illustrate the dangers of treating disaggregated units such as segments and zip codes as independent. Indeed, as highlighted above, a perfectly segmented economy would not be able to distinguish the spillover effects of more versus less integrated neighborhoods. It would also provide misleading summary information about the demand for housing. For example, at our estimated parameters, mean popularity is 1.03 for Daly City and 0.75 for the Outer Mission, which features more broad searchers. In contrast, in the perfectly segmented benchmark economy introduced in Section V, mean popularities are 1.008 and 1.014 for Daly City and the Outer Mission, respectively. An econometrician using a perfectly segmented economy to assess the effects of new construction would therefore incorrectly expect similarly sized increases in local inventory in response to construction in the Outer Mission and Daly City. In addition, the econometrician would predict no spillovers to other segments for construction in either market.

VI. Conclusions and Segmented Search in Other Markets

Most search markets feature competition between broad and narrow searchers. We show that observing the structure of these search clienteles is important for understanding the forces behind equilibrium market outcomes such as the shape of the Beveridge curve, as well as the response of different market segments to shocks. We also demonstrate how data from online search behavior can allow researchers to overcome the challenge of measuring clientele patterns. We expect that similar data from online services such as Facebook, LinkedIn, Tinder, ZipRecruiter, and Indeed will allow researchers to measure the clientele structure in other search markets, from dating to job search. This will improve our understanding of both the cross-sectional patterns across submarkets, as well as the response of these markets to shocks.

While our analysis highlights the importance of understanding the interaction of broad and narrow searchers in the housing market, our insights are likely to also be important in other search markets. For example, Treasury securities are sold in over-the-counter search markets. In these markets, some investors might be particularly interested in purchasing Treasuries of certain maturities (e.g., pensions funds engaging in duration matching might only buy long-duration Treasuries), but there might be other buyers, such as hedge funds, that consider a broader range of maturities. Understanding the segmentation of the buyer clientele, and the interaction of broad and narrow investors at different maturities, is important for determining the optimal maturity structure of newly issued debt. For example, monetary policy interventions, such as the maturity extension program (Operation Twist), which aimed at flattening the yield curve by selling short-maturity Treasuries and buying long-maturity Treasuries, will be less effective in the presence of "broad" investors who are indifferent between buying Treasuries with a wide range of maturities (see Swanson 2011, Greenwood and Vayanos 2014).

Similarly, the degree of segmentation in labor market search is substantial and time-varying. For example, recent research has documented that job seekers in areas with depressed housing markets apply for fewer jobs that require relocation, because of the difficulties with selling underwater homes (see Brown and Matsa 2019). Our model allows researchers to understand the extent to which the resulting increase in regional segmentation of labor search contributes to a decline in the ability of labor flows to facilitate regional risk sharing. In addition, labor market segmentation across industry and occupational groups has increased as a result of the specialization of human capital, with many vacancies only attracting applications from a small set of highly specialized job seekers. The mechanisms highlighted in this paper show how such changes in segmentation influence the labor market effects of immigration, and affect the efficacy of policies such as targeted job training programs (e.g., Peri and Sparber 2009).

The dating market provides a further setting in which some searchers with very narrow preferences interact with other searchers who are less particular about the characteristics of their preferred match. More recently, the rise of online dating services has increased the ability of narrow searchers to target their search to their particular preferences. An interesting question is whether the resulting increase in segmentation of the dating market has contributed to longer "times on market" and the increase in the age of marriage.

REFERENCES

Albrecht, James, Axel Anderson, Eric Smith, and Susan Vroman. 2007. "Opportunistic Matching in the Housing Market." *International Economic Review* 48 (2): 641–64.

Anenberg, Elliot, and Patrick Bayer. 2014. "Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem." NBER Working Paper 18980.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel. 2018a. "The Economic Effects of Social Networks: Evidence from the Housing Market." *Journal of Political Economy* 126 (6): 2224–76.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2018b. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives* 32 (3): 259–80.

Bailey, Michael, Patrick Farrell, Theresa Kuchler, and Johannes Stroebel. 2019a. "Social Connectedness in Urban Areas." NBER Working Paper 26029.

Bailey, Michael, Drew M. Johnston, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2019b. "Peer Effects in Product Adoption." NBER Working Paper 25843.

Bailey, Michael, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel. 2019c. "House Price Beliefs and Mortgage Leverage Choice." *Review of Economic Studies* 86 (6): 2403–52.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588–638.

Brown, Jennifer, and David A. Matsa. 2019. "Locked in by Leverage: Job Search during the Housing Crisis." https://www.sciencedirect.com/science/article/abs/pii/S0304405X19302697.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 2016. "Understanding Booms and Busts in Housing Markets." *Journal of Political Economy* 124 (4): 1088–1147.

Davis, Morris A., and Stijn Van Nieuwerburgh. 2015. "Housing, Finance, and the Macroeconomy." In *Handbook of Regional and Urban Economics,* Vol. 5, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, 753–811. Amsterdam: North-Holland.

Díaz, Antonia, and Belén Jerez. 2013. "House Prices, Sales, and Time on the Market: A Search-Theoretic Framework." *International Economic Review* 54 (3): 837–72.

Einav, Liran, Theresa Kuchler, Jonathan Levin, and Neel Sundaresan. 2015. "Assessing Sale Strategies in Online Markets Using Matched Listings." *American Economic Journal: Microeconomics* 7 (2): 215–47.

Genesove, David, and Lu Han. 2012. "Search and Matching in the Housing Market." *Journal of Urban Economics* 72 (1): 31–45.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Andreas Weber. 2015. "Climate Change and Long-Run Discount Rates: Evidence from Real Estate." NBER Working Paper 21767.

Goodman, Allen C., and Thomas G. Thibodeau. 1998. "Housing Market Segmentation." *Journal of Housing Economics* 7 (2): 121–43.

Greenwood, Robin, and Dimitri Vayanos. 2014. "Bond Supply and Excess Bond Returns." *Review of Financial Studies* 27 (3): 663–713.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst. 2013. "Endogenous Gentrification and Housing Price Dynamics." *Journal of Public Economics* 100: 45–60.
Guren, Adam M. 2018. "House Price Momentum and Strategic Complementarity." *Journal of Political Economy* 126 (3): 1172–1218.
Guren, Adam M., and Timothy J. McQuade. 2014. "How Do Foreclosures Exacerbate Housing Downturns?" Unpublished.
Halket, Jonathan, and Matteo Pignatti Morano de Custoza. 2015. "Homeownership and the Scarcity of Rentals." *Journal of Monetary Economics* 76: 107–23.
Han, Lu, and William C. Strange. 2015. "The Microstructure of Housing Markets: Search, Bargaining, and Brokerage." In *Handbook of Regional and Urban Economics*, Vol. 5, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, 813–86. Amsterdam: North-Holland.
Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun. 2014. "Search and the Dynamics of House Prices and Construction." *American Economic Review* 104 (4): 1172–1210.
Islam, Kazi Saiful, and Yasushi Asami. 2009. "Housing Market Segmentation: A Review." *Review of Urban & Regional Development Studies* 21 (2–3): 93–109.
Krainer, John. 2001. "A Theory of Liquidity in Residential Real Estate Markets." *Journal of Urban Economics* 49 (1): 32–53.
Kurlat, Pablo, and Johannes Stroebel. 2015. "Testing for Information Asymmetries in Real Estate Markets." *Review of Financial Studies* 28 (8): 2429–61.
Landvoigt, Tim, Monika Piazzesi, and Martin Schneider. 2015. "The Housing Market(s) of San Diego." *American Economic Review* 105 (4): 1371–1407.
Leishman, Chris. 2001. "House Building and Product Differentiation: An Hedonic Price Approach." *Journal of Housing and the Built Environment* 16 (2): 131–52.
Manning, Alan, and Barbara Petrongolo. 2017. "How Local Are Labor Markets? Evidence from a Spatial Job Search Model." *American Economic Review* 107 (10): 2877–2907.
National Association of Realtors. 2011. "Profile of Home Buyers and Sellers 2011." Technical Report. Washington, DC: NAR.
National Association of Realtors. 2013. "The Digital House Hunt." Technical Report. Washington, DC: NAR.
Ngai, L. Rachel, and Silvana Tenreyro. 2014. "Hot and Cold Seasons in the Housing Market." *American Economic Review* 104 (12): 3991–4026.
Ortalo-Magné, François, and Sven Rady. 2006. "Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints." *Review of Economic Studies* 73 (2): 459–85.
Peri, Giovanni, and Chad Sparber. 2009. "Task Specialization, Immigration, and Wages." *American Economic Journal: Applied Economics* 1 (3): 135–69.
Piazzesi, Monika, and Martin Schneider. 2009. "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." *American Economic Review* 99 (2): 406–11.
Piazzesi, Monika, and Martin Schneider. 2017. "Housing and Macroeconomics." In *Handbook of Macroeconomics,* Vol. 2, edited by John B. Taylor and Harald Uhlig, 1547–1640. Amsterdam: North-Holland.
Poterba, James M. 1991. "House Price Dynamics: The Role of Tax Policy and Demography." *Brookings Papers on Economic Activity* (2): 143–203.
Stroebel, Johannes. 2016. "Asymmetric Information about Collateral Values." *Journal of Finance* 71 (3): 1071–1112.
Swanson, Eric T. 2011. "Let's Twist Again: A High-Frequency Event-Study Analysis of Operation Twist and Its Implications for QE2." *Brookings Papers on Economic Activity* (1): 151–88.
Trulia. 2013. "The Largest Audience of Real Estate Consumers." Technical Report. New York: Trulia.
Wheaton, William C. 1990. "Vacancy, Search, and Prices in a Housing Market Matching Model." *Journal of Political Economy* 98 (6): 1270–92.