# Measuring Housing Quality Using Revealed Preference: A Geographic PageRank Approach *

Alex Bell † Sophie Calder-Wang‡ Shusheng Zhong§

*Extended Abstract*

March 13, 2025

**Abstract**

This paper introduces Geographic PageRank (GPR), an innovative measure of locational desirability that leverages migration decisions to assess housing quality at the neighborhood level. Conceptually, GPR employs a recursive algorithm that uses the full network of migration flows to derive a micro-founded measure of place quality. Our GPR measure can be particularly useful in settings where unobserved housing quality has led to biased, or even wrong-signed, estimates for standard hedonic regressions of housing prices on amenities. We also demonstrate how the GPR framework can be extended to finer geographic units and adapted to capture temporal dynamics, providing a versatile tool for housing analysis.

*Keywords:* PageRank, Housing and Neighborhood Quality, Amenities *JEL Code*: R32, R31, C39

*This paper benefited from conversations with seminar and conference participants at the University of Toronto, SUNY Albany, AEA, AREUEA, SEA, and UEA. Calder-Wang thanks the Zell/Lurie Real Estate Center at Wharton for financial support.

†Department of Economics, Georgia State University *ambell@gsu.edu*

‡The Wharton School, University of Pennsylvania *sophiecw@wharton.upenn.edu*

§Kellogg School of Management, Northwestern University *shusheng.zhong@kellogg.northwestern.edu*

# 1 Introduction

Understanding housing quality is fundamental for urban policy, real estate valuation, and broader economic analysis. Yet, accurately measuring housing quality has remained elusive, where unobserved housing quality can bias estimates, making it challenging to accurately assess the prices of residential and neighborhood attributes.

This paper introduces Geographic PageRank (GPR), a novel metric that leverages the network of migration flows as a revealed-preference mechanism to measure the quality rankings of housing, especially at the neighborhood level. We begin with the well-known PageRank algorithm from computer science developed in Page (2001), which had been the foundational piece underlying much of Google's search rankings. Instead of using the algorithm to model the network connections of webpages, we adapt the algorithm to model the network connection of places as characterized by the migration flows that connect different places into a network. Intuitively, the algorithm delivers a measure of the desirability of a particular location by computing it recursively: locations that draw in migration are considered desirable, whereas locations that draw in migration from other desirable locations are considered even more desirable. Conversely, an undesirable location is characterized by both out-migration and losing people to other undesirable locations. In doing so, the GPR measure effectively encapsulates information embedded in the entire migration network, rather than conventional metrics such as net migration, which is defined based on the net flow of the focal location itself.

The PageRank algorithm's ability to characterize information from the entire network using a recursive method has proven to be very effective across many domains. The original PageRank algorithm, developed by Larry Page, was extremely effective in producing informative rankings and was foundational to Google's success in becoming the most popular search engine. In computer science, the use of PageRank type of algorithm has become pervasive and standard in many network problems (Berkhin, 2005), but its uses have been extended to many other fields ranging including biology, chemistry, neuroscience, and physics (Gleich, 2015).

In economics, using PageRank to characterize network properties as a measure of economic features is emergent. In labor economics, a seminal paper is Sorkin (2018), which uses worker's employer-to-employer transitions history to characterize the network of firms and, therefore, a ranking of firm quality. The core intuition is, excluding layoffs, when workers switch firms, they, on average, move from worse to better firms, as revealed by their preferences. Such measures of firm quality then become a crucial input to demonstrating the role of compensating differentials in labor markets. Related work in the labor literature has shown the effectiveness of migration-based measures in capturing firm quality

in a variety of contexts (Bagger and Lentz, 2019; Morchio and Moser, 2020; Taber and Vejlin, 2020).

In urban economics, however, there has been no systematic effort in any published work to leverage PageRank as a measure of housing quality.[1] Recent developments in the housing literature, such as Epple, Quintero and Sieg (2020) and Landvoigt, Piazzesi and Schneider (2014) all highlight the vertical measure of housing quality as the primary dimension where housing markets become segmented and differentiated. In this regard, the Geographic PageRank can be used to capture the vertical dimension of the housing market, namely, housing quality. Analogous to the labor context, the core intuition is that, unless forced to, households, on average, move from worse to better housing arrangements, as revealed by their preferences.

Concretely, we start with county-to-county migration data from the Internal Revenue Service (IRS) to produce the Geographic PageRank for all U.S. counties. We show that, while it correlates with net migration rates, substantial differences remain. In addition, we also apply the GPR framework to finer spatial units. For example, we compute measures of locational desirability at more granular geographic units (ZCTA) using individual-level migration data from Infutor. We can also do this using data from Data Axle (formerly Info Group), which allows us to focus on the migration flow of working-age households. In general, the GPR measure of housing quality can be readily extended to various levels of geographic granularity, ranging from very granular census tract-level rankings to county-level, city-level, or even metropolitan-level rankings. Lastly, we can also construct GPR measures for different time periods, capturing changes in the desirability of different places.

To illustrate how GPR can correct biases in standard hedonic models, we cite one existing use case on evaluating the implicit price of environmental amenities—specifically air quality—highlighting how GPR can serve as an effective anti-instrument to recover sensible estimates of the capitalization effect of air pollution on housing markets. See Bell et al. (2024) for more details of the application and the econometric framework. Notably, GPR has allowed us to produce the implicit amenity prices using observational data alone, yet it is able to produce estimates that are largely in line with Chay and Greenstone (2005), which offers the best evidence in the literature that relies on natural experiments.

In sum, a comprehensive Geographic PageRank measure can contribute to better measurement of housing quality at various levels of geographic granularity, as well as provide a consistent benchmark over time to evaluate the relative quality of different places as they change. Besides providing useful

[1] Unpublished work by Fogel (2021) applies a similar approach to estimate city values based on migration flows. His work focuses on city-to-city movements, with an emphasis on understanding its relationship with local labor demand shocks, which is found to be null. By contrast, our GPR measures focus on extract the set of migration-flows that can be best fitted for measuring housing quality with greater geographic granularity, and thus becomes useful for studying local housing markets and the capitalization of amenities at the neighborhood level.

descriptive results, such as a better-measured housing quality, GPR has also been shown to improve estimates of implicit prices of housing amenities by providing better controls for unobserved components. Lastly, by providing a better measure of housing quality, we also expect use cases for GRP as it becomes included in structural models of housing and neighborhood choices and potentially many more beyond.

The remainder of the paper is organized as follows. Section 2 provides the mathematical foundations for the PageRank algorithm. Section 3 describes various sources of migration data used. Section 4 describes the rankings produced at the county and census tract levels. It also describes how GPR can be extended in various ways. Section 5 cites an example of using GPR to price environmental amenities in the housing market. Section 6 concludes.

## 2 Mathematical Foundations of Geographic PageRank

In this section, we describe the mathematical foundations to the PageRank algorithm.

The key idea behind "Geographic PageRank" is to capture a ranking of geographic locations using a recursive logic based on migration: by revealed preference, when households move, they must on average move to a "better" place than their previous place, and places that draw people from "better" places are considered "best" places. As such, we use a matrix of migration to arrive at a ranking of places based on such recursive definition of common desirability.

At a conceptual level, our proposed Geographic PageRank algorithm ranks a location by its share in the stationary distribution of household locations in accordance to the migration matrix. Specifically, let the migration matrix $M_{i,j}$ represent the fraction of population migrating from location $j$ to $i$. In other words, the matrix specifies the destination locations for everyone from location $j$, and $M_{i,j}$ represents a transition matrix (in a column format) such that every column sums to one

$$\forall j : \sum_i M_{i,j} = 1. \tag{2.1}$$

In its simplest form, the PageRank algorithm computes the column vector $v$ such that

$$Mv = v. \tag{2.2}$$

Equivalently, $v$ represents the eigenvector associated with the transition matrix $M$.[2] In this sense, the resulting eigenvector $v$ represents the stationary distribution of the population if they migrated according to the transition matrix $M$ indefinitely: namely, the transition matrix $M$ on the distribution $v$ recovers $v$. As a last step, we obtain the ordinal "rank" of location $i$ based on the rankings of the value of each element $v_i$.

To provide some intuition behind the PageRank algorithm, consider the following stylized example. In Panel (A) of Figure 1, there are three locations A, B, and C. Assume that they are equal-sized cities and we observe that 10% of C's residents migrated to A and another 10% migrated to B. Performing the PageRank algorithm on the resulting migration matrix $M$ results an equal ranking between A and B, which are both ranked above location C.

In Panel (B) of Figure 1, we also consider three locations A, B, and C. However, in this case, we find 5% of C's residents migrated to A, and 15% migrated to B. Moreover, we also find 5% of B's residents also migrated to A. In this case, the PageRank algorithm ranks A above B, and then above C. Notably, the PageRank algorithm contains more information than simple counts of net-migration: even though both location A and B see a net-in migration of 10% in both examples, the PageRank algorithm ranks A above B in the second example, as A is drawing residents from B, which is in turn more desirable than C.

Beyond its widespread application in computer science, Sorkin (2018) is the first to use the PageRank algorithm in economics and it is applied to ranking a set of U.S. firms in the labor market context. Indeed, Sorkin (2018) proceeds to use the ranking of firms as a way to measure how much wage dispersion is due to job amenities. We highlight that there is a natural parallel between the challenges of the empirical valuation of job amenities and the empirical valuation of neighborhood amenities, even though there has not been much overlap between the two literature in recent works. As such, our work fits into this important theoretical and methodological gap.

## 3 Data

In this section, we describe the various data sources that we have access to that we use, or plan to use, to measure geographic mobility, which serves as the basis for the PageRank algorithm.

[2] In practice, the PageRank algorithm also allows for a damping factor $d$ such that every period a small fraction $1-d$ from each node will be randomly redistributed to all other nodes to ensure numerical stability. We find that the ordinal rank of the nodes are not sensitive to the choice of the damping factor. As such, the PageRank algorithm could be viewed as one of the spectral methods derived from eigenvector centrality in network analysis, but has properties to ensure desirable numerical stability, especially with disconnected graphs.

### 3.1 IRS County-to-County Migration

The first data source is the county-to-county migration data from the Internal Revenue Service (IRS) Statistics of Income (SOI) Division, which is "based on year-to-year address changes reported on individual income tax returns filed with the IRS."[3] The data is publicly available and covers the entire universe of population with Forms 1040 filings. Another advantage of this data is that it has a long time coverage, from 1991 to present. The main caveats of the data are that it does not cover migration at finer geographic level or by demographics, and it does not include non-filers. Furthermore, flows between counties with fewer than 20 returns are suppressed for confidentiality protection.

### 3.2 Infutor

To measure geographic mobility at a finer resolution, we also use the Infutor data. The Infutor data contain individual-level mobility data across the United States that records the past ten residences at the street address level for each individual. Coverage is thought to be broad, though variable over time. For illustration purpose, we use a subsample of the data that contains all the individuals who have ever lived in the state of California to measure migration flows between California ZIP Code Tabulation Areas (ZCTA) in the year 2019.

### 3.3 Data Axle

While IRS and Infutor data can both produce migration counts, they are rather limited in terms of other demographic information associated with the households. To the extent that researchers may be interested in migration choices made by certain types of households (e.g., those under 65), we plan to use the panel of household address histories and associated demographic features available in Data Axle (formerly Info Group) to produce better GPR measures that best fit the revealed preference setting.

### 3.4 ACS

Finally, we also plan to use the migration information from the American Community Survey (ACS) micro-data. ACS surveys its respondents in terms of their home location, measured in Public Use Microdata Area (PUMA), one year ago, as well as questions about the reasons for the move (e.g., housing related, work related, family related etc.) More importantly, while the publicly available ACS micro data has coarse geographic information (only in terms of previous PUMA, which still averages about 100,000 in population) compared to Data Axle, it contains a much wider set of household

[3] https://www.irs.gov/pub/irs-soi/1819inpublicmigdoc.pdf

demographics, far richer than any alternative sources, thus, allowing us to narrow down the set of migration links that are most appropriate to the revealed preference setting.

# 4 Geographic PageRank Estimates

## 4.1 A Ranking of U.S. Counties

To implement the Geographic PageRank algorithm at the national level, we use the outflow version of the IRS 2018-2019 county-to-county Migration Data, which records "the number of residents leaving a State or county and where they went." Table 1 shows the summary statistics for the IRS migration data. 2,824 out of the total 3,033 counties has more than 20 migration flows to another county. Within these counties, the mean total outflow is 2090.6 with a standard deviation of 6802.9, and the median total outflow is 348. There are 47,974 origin-destination county pairs out of all the 9,196,056 possible pairs with more than 20 migration flows. Within these county pairs, the mean migration flow is 123.1 with a standard deviation of 468.9, and the median migration flow is 41.

Next, we obtain the corresponding rankings of all U.S. counties using the Geographic PageRank algorithm described above.[4] Figure 2 illustrate the rankings on the map: green shades indicate higher-ranked (i.e., more desirable) counties and red shades indicate lower-ranked (i.e., less desirable) counties. We generally find higher-ranked counties along the coastal regions. Table 2 shows the top 20 counties by its Geographic PageRank. The top 10 cities are Phoenix, Houston, Los Angeles, Dallas, Chicago, Fort Worth, Seattle, Austin, Minneapolis, and San Antonio. We observe that many of them are in the Sun Belt, reflecting recent migratory trends.

In general, places that see population growth are on average ranked higher. The top panel of Figure 4 shows the binscatter between Geographic PageRank and net migration, clearly indicating a positive relationship. However, there remains significant dispersion for the same level of net population change, and the average relationship is also non-linear, as shown in the scatter plot in the bottom panel of Figure 4. The remaining dispersion suggests that the recursive nature of the PageRank algorithm synthesizes more information from the full migration matrix $M$, as opposed to the simple summation operation that computes the net migration alone.

## 4.2 More Granular Rankings from Infutor Data

Although considerably noisier at present, as a proof of concept, our dataset of residential mobility from Infutor also allows us to rank finer-grained geographies. Figure 3 shows a map of Geographic

[4] Our specific implementation assumes a damping factor of $d = 0.85$, which corresponds roughly to the average stay rate.

PageRank for California's ZCTA's, based on moves recorded by Infutor in 2019. In this version of the paper, we do not show additional results leveraging ZCTA-level moves due to the increased noise associated with small sample sizes at finer-grained geographies. However, we are in the process of cleaning and incorporating additional years and geographic coverage of the data, which will allow for more reliable estimates to study more granular mobility patterns.

### 4.3 A Comprehensive Set of Measurements (under Development)

We also briefly outline a number of key next steps that we plan on taking to produce a more comprehensive set of rankings to be used for various purposes. We anticipate to make all the rankings that are derived from public data also readily available for research use.

1. Ranking by geography
    - Rankings of counties, ZCTAs, and census tracts for all of the U.S.
2. Ranking over time
    - Yearly rankings of counties from 1991 to present (using IRS migration data)
    - 5-year rankings of ZCTAs and census tracts from 2010 to present (using Data Axle migration data)
3. Ranking by demographic groups
    - Yearly rankings of counties from 2006 to present (using ACS Micro data)
        - Subgroups by age (e.g., over 65), labor market status, and stated reason for moving

It is worth noting that the ranking by demographic groups are not only of interest in itself, but different measures may be more appropriate for the applications in hand. For example, there are reasons to suspect that migratory tendencies may change as households age because their preferences for amenities change. This limitation must be grappled with if one wishes to apply Geographic PageRank to pricing weather as an amenity, which Glaeser, Kolko and Saiz (2001) find to be among the most important amenities in driving home price differentials. For this reason, we will leverage the panel available in the Data Axle data to investigate the extent to which rankings of places are stable or change over individuals' life cycles by computing ranks based on specific age cohorts over time. However, our preliminary analysis has shown that the rankings are incredibly similar even among different age groups.

In addition, given the richness of household demographics available in the ACS data, we can also compute measures of Geographic Page rank that are based on certain types of migration flows that are most fitting for the revealed-preference framework, for example, by focusing those who are in the labor market and who are moving for housing related reasons.

# 5 Application: Using GPR to Value Air Quality

In this section, we cite an application from Bell et al. (2024) where we illustrate that the Geographic PageRank can be used as a measure of unobserved housing quality to obtain sensible estimates of the price of air pollution. We demonstrate how it works in both the general non-linear case and the more specific linear case. We show that the anti-IV-based estimates are able to address the "wrong-signed" problems that seem to be pervasive in the estimation of similar environmental amenities. We also find that the estimates are in-line with other existing reduced-form estimates using natural experiments.

First, we obtain a correctly-signed negative price for air pollution in the general setting without assuming linearity of the index mapping function, as shown in Figure 5. Specifically, the estimation procedure goes as follows: First, we divide both price and amenities into quantiles, and in this case, we have divided them each into $N = 11$ equally sized quantiles. Then, the Cartesian product of each price quantile and each air pollution quantile generates a total of $N^2$ cells. For each cell, we collect all households who belong to that cell and compute the average Geographic PageRank of members of the cell, which captures the predicted GPR conditional on price and amenity.

The pattern in Figure 5 indicates a trade-off between home price and air quality. It shows a clear increase in the predicted Geographic PageRank in the diagonal direction: the best-ranked places (i.e. dark green) are in the upper-right hand corner and the worst-ranked places (i.e. dark red) are in the bottom-left corner. As such, the upward-sloping blue line in Figure 5 indicates the direction of expanding choice sets, whereas the perpendicular direction (i.e., the dark green line) shows the level sets: the set of cells all with the same average PageRank but with varying combinations of price and amenity. In other words, the direction perpendicular to the direction of choice-set expansion shows the implied trade-off between home price and clean air: places with cleaner air are priced higher than places with more polluted air, conditioning on being in the same housing quality segment, which are proxied by the predicted Geographic PageRank.

Next, if we take on the lens of linearity, then we can use the "ratio-of-coefficients" method to obtain a single coefficient estimate of the amenity price. To visually benchmark the magnitude of our estimate in a linear model, Figure 6 compares the estimated price elasticity with known estimates in

the literature, specifically, with Chay and Greenstone (2005). The shaded region indicates plausible point estimates of the housing price elasticity with respect to total suspended particulates (TSPs), which ranged from -0.20 to -0.35. The first three dots in Figure 6 shows that the conventional OLS or OLS with controls appear to generate wrong-signed and downward-biased estimates. The last dot in Figure 6 shows a significantly negative elasticity of -.40 estimated by our anti-IV-based approach, and the confidence interval overlaps with the range of point estimates in Chay and Greenstone (2005). Though it should not necessarily be expected ex ante that they are the same given the nuances in how price is conceptualized, the fact that our estimate of the price of air pollution is indistinguishable from the range of estimates in the quasi-experimental literature lends some support to the credibility of our approach.

# 6 Conclusion

This paper introduces Geographic PageRank (GPR) as a novel, network-based measure of housing and neighborhood quality, leveraging migration flows to capture revealed preferences for locational desirability. By adapting the recursive logic of the PageRank algorithm, GPR synthesizes information from the full migration network, providing a more robust measure of housing quality than conventional approaches based on net migration or hedonic regressions.

Since GPR can be measured at multiple geographic scales, this flexibility makes it particularly valuable for applications in urban and environmental economics, where housing quality and amenity values vary at a highly localized level. Furthermore, given the measure is inherently noisy, we show how single latent index models are particularly important to unlock the full potential of this network-based measure of housing quality (Bell, 2022; Bell et al., 2024).

We also plan to expand from the existing set of results by incorporating demographic heterogeneity in migration flows, exploring dynamic changes in place quality over time, and integrating GPR into structural models of household location choice. As migration patterns continue to evolve, tools like Geographic PageRank can offer valuable insights into urban dynamics, housing markets, and spatial inequality.

## References


**Bagger, Jesper, and Rasmus Lentz.** 2019. "An Empirical Model of Wage Dispersion with Sorting." *The Review of Economic Studies*, 86(1): 153–190.

**Bell, Alex.** 2022. "Job Amenities and Earnings Inequality." Available at SSRN: http://dx.doi.org/10.2139/ssrn.4173522.

**Bell, Alex, Stephen B. Billings, Sophie Calder-Wang, and Shusheng Zhong.** 2024. "An Anti-IV Approach for Pricing Residential Amenities: Applications to Flood Risk *."

**Berkhin, Pavel.** 2005. "A Survey on PageRank Computing." *Internet Mathematics*, 2(1): 73–120.

**Chay, Kenneth Y., and Michael Greenstone.** 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, 113(2): 376–424. Publisher: The University of Chicago Press.

**Epple, Dennis, Luis Quintero, and Holger Sieg.** 2020. "A New Approach to Estimating Equilibrium Models for Metropolitan Housing Markets." *Journal of Political Economy*, 128(3): 948–983. Publisher: The University of Chicago Press.

**Fogel, Jamie.** 2021. "Valuing American Cities: A Revealed Preference Approach."

**Glaeser, Edward L., Jed Kolko, and Albert Saiz.** 2001. "Consumer city." *Journal of Economic Geography*, 1(1): 27–50.

**Gleich, David F.** 2015. "PageRank Beyond the Web." *SIAM Review*, 57(3): 321–363.

**Landvoigt, Tim, Monika Piazzesi, and Martin Schneider.** 2014. "Housing Assignment with Restrictions: Theory and Evidence from Stanford University's Campus." *American Economic Review*, 104(5): 67–72.

**Morchio, Iacopo, and Christian Moser.** 2020. "The Gender Pay Gap: Micro Sources and Macro Consequences." Social Science Research Network SSRN Scholarly Paper ID 3176868, Rochester, NY.

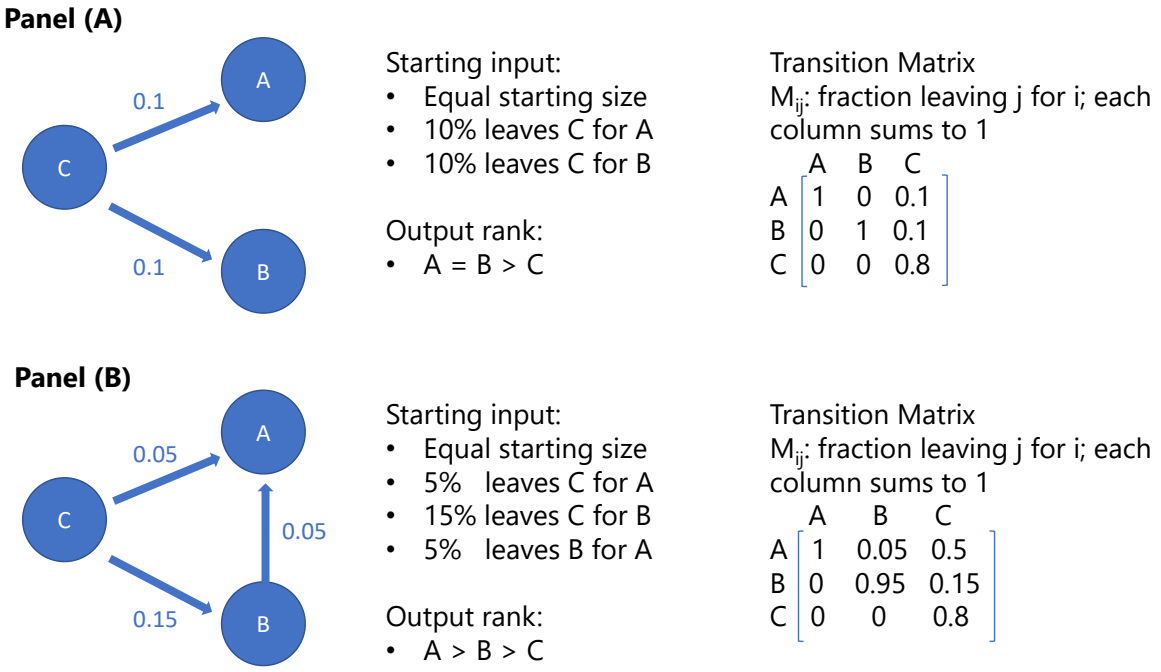
**Page, Lawrence.** 2001. "Method for node ranking in a linked database."

**Sorkin, Isaac.** 2018. "Ranking Firms Using Revealed Preference*." *The Quarterly Journal of Economics*, 133(3): 1331–1393.

**Taber, Christopher, and Rune Vejlin.** 2020. "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market." *Econometrica*, 88(3): 1031–1069. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14441.

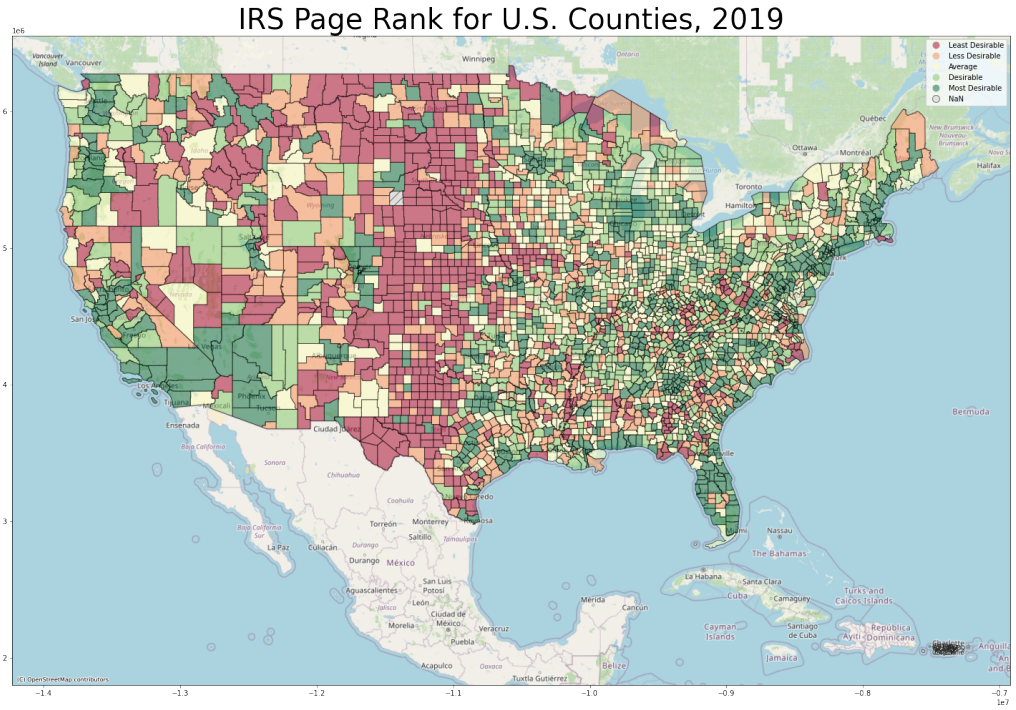
# 7 Figures

Figure 1: Example of PageRank Algorithm



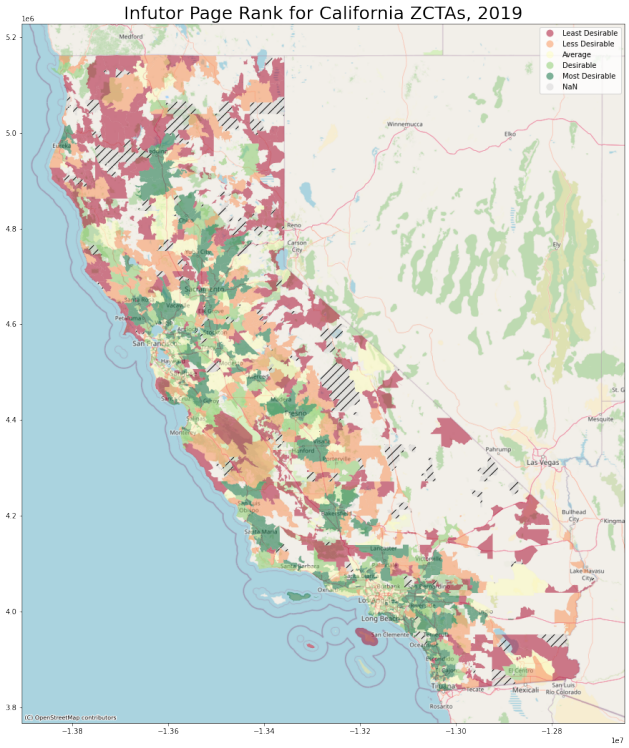

*Notes:* The figure above illustrates a simple example of the PageRank algorithm. In the top panel (A), there are three locations A, B, and C, where there is a net migration from location C to both A and B in equal proportions. The adjacency matrix $M_{i,j}$ indicates the transition probabilities in the column format, namely, $M_{i,j}$ represents the fraction departing from node $j$ for node $i$. As such, the normalization requires $\forall i : \sum_j M_{i,j} = 1$. The resulting rank is $A = B > C$. By contrast, the bottom panel (B) illustrates another example where there is not only a migration from C, but also some migration from B to A. Note that the net in-migration for both examples for location A and B are the same, namely, a net migration of 10%. However, because location A is drawing people from the more desirable location B, as opposed to C, the resulting rank has A being ranked above B, namely $A > B > C$.

Figure 2: County-level PageRank using IRS County Migration Data



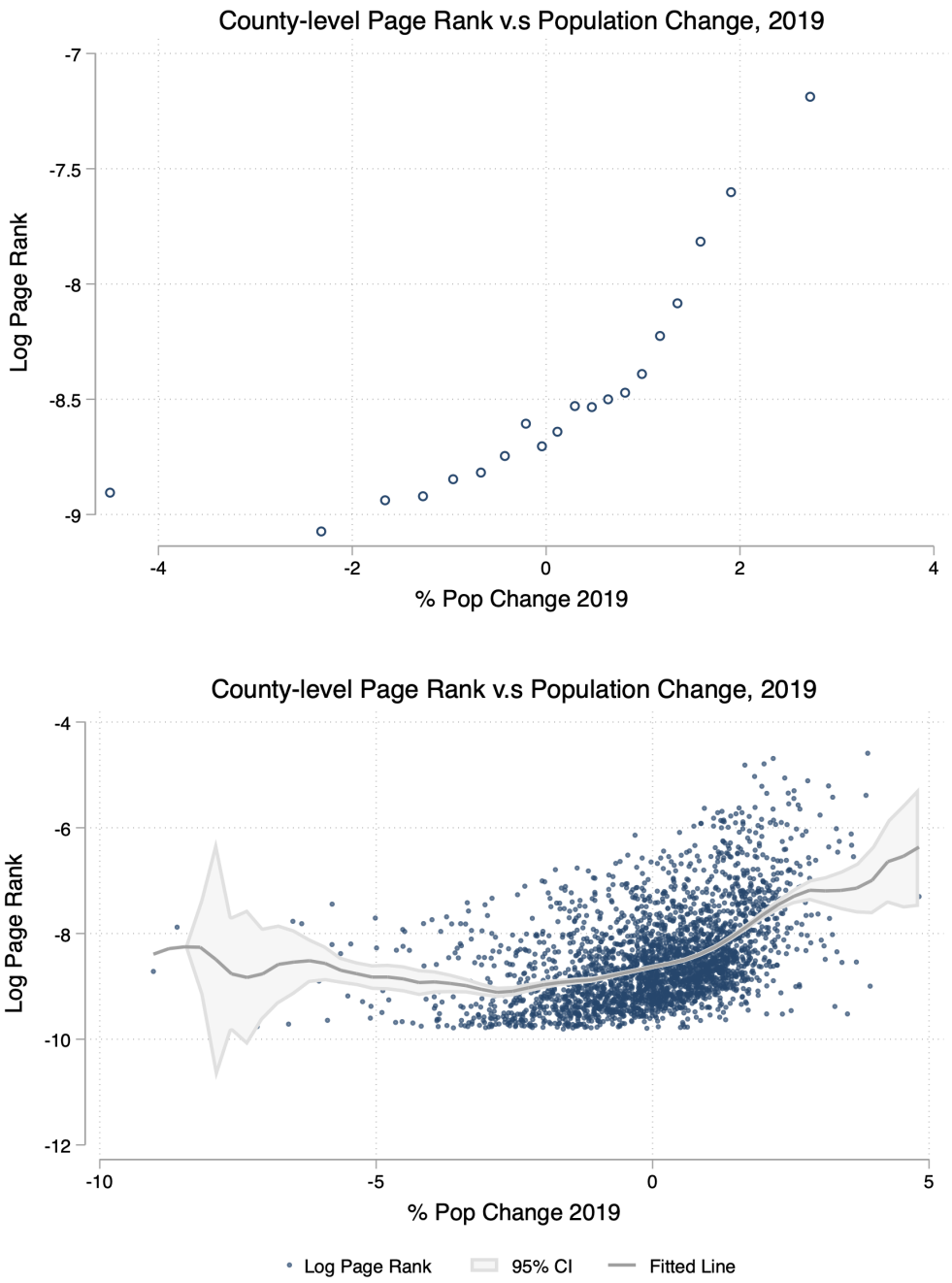

*Notes*: This map shows geographic PageRank at the county-level constructed using nationwide IRS county migration flow data in 2019. Green shades indicate higher-ranked (i.e., more desirable) counties and red shades indicate lower-ranked (i.e., less desirable) counties.

Figure 3: ZCTA-level PageRank using Infutor Migration Data in California



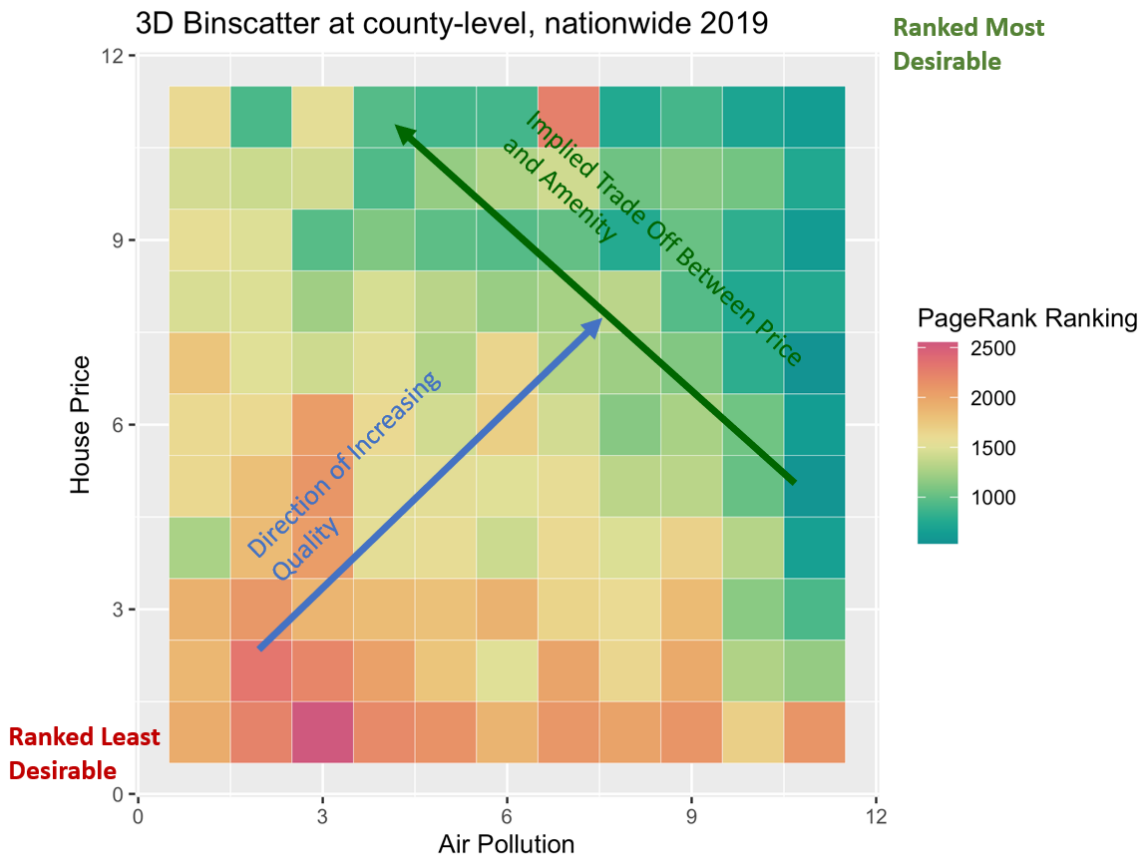

*Notes*: This map shows geographic PageRank at the ZCTA-level constructed using Infutor data in 2019. Green shades indicate higher-ranked (i.e., more desirable) ZCTAs and red shades indicate lower-ranked (i.e., less desirable) ZCTAs.

Figure 4: Geographic PageRank v.s. Net Migration



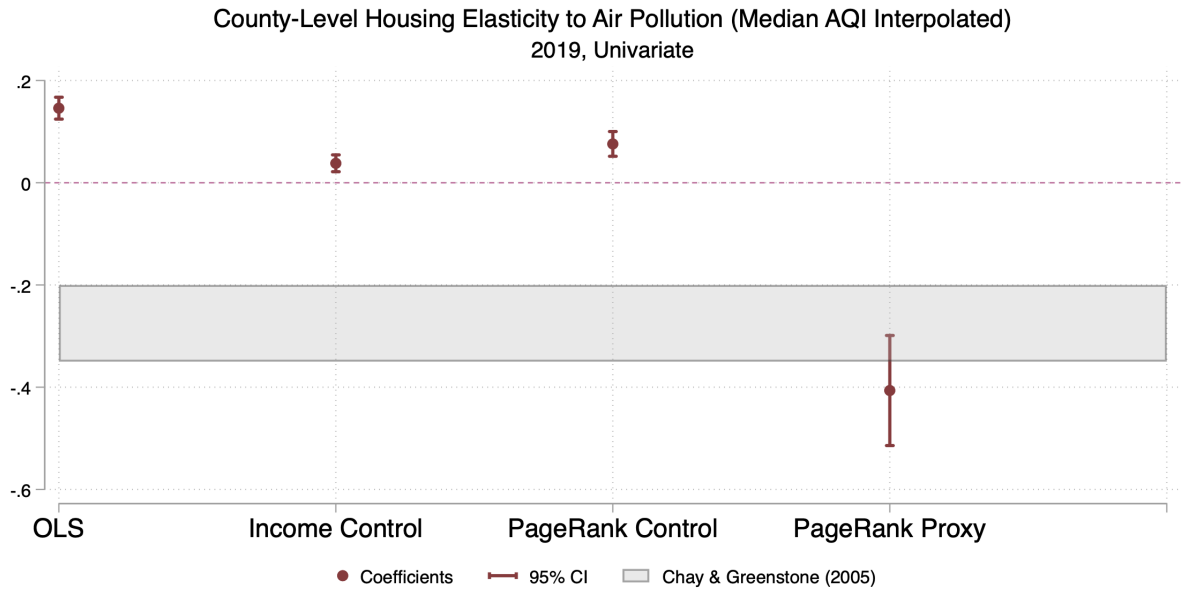

*Notes*: The binscatter (top panel) and the scatter plot (bottom panel) shows the relationship between Geographic PageRank and net migration. The y-axis represents the logarithm of the elements of the PageRank eigenvector. In general, places that see population growth are ranked higher. However, there remains significant dispersion for the same level of net migration and the average relationship is also non-linear.

Figure 5: 3D Binscatter of Predicted PageRank by Price and Air Pollution




*Notes*: This figure categorized the counties into 121 (11x11) cells. Each cell represents counties that lie within certain house price decile x air pollution decile. For example, the upper-right-most cell represents counties in top income quantile and top air pollution quantile. The color of each cell represents the median Geographic PageRank Ranking of all the counties in this cell. The PageRank algorithm ranks the most desired place as number 1. Green shades represent higher-ranked (i.e., more desirable) places and red shades represent lower-ranked (i.e., less desirable) places.

Figure 6: Pricing Air Pollution Using Anti-IV Approach




*Notes*: The estimated coefficients on this figures correspond to the estimated price elasticity to median AQI from regression results in Table **??**.

# 8 Tables

Table 1: Summary Statistics for 2018-2019 IRS County-to-County Migration Data

| | County | | Origin-Destination Pairs |
|---|---|---|---|
| | Outflows | Inflows | Migration Flows |
| | (1) | (2) | (3) |
| Mean | 2090.62 | 2142.20 | 123.06 |
| SD | 6802.89 | 6415.80 | 468.86 |
| Min | 20 | 20 | 20 |
| 25th Percentile | 126.5 | 128.5 | 27 |
| Median | 348 | 362.5 | 41 |
| 75th Percentile | 1130.5 | 1211 | 81 |
| Max | 145204 | 111327 | 20908 |
| N | 2824 | 2756 | 47974 |

*Notes*: Migration flows between counties with less than 20 counts are suppressed in the data. Column (1) reports the county-level summary statistics for the total migration flows out of each county into other counties, and Column (2) reports the county-level summary statistics for the total migration flow into each county from other counties. Column (3) reports the summary statistics for migration flows at the level of origin county-destination pairs.

Table 2: Top Ranked Counties in the U.S. (2019)

| FIPS | Ranking | County Name | MSA Name |
|---|---|---|---|
| 4013 | 1 | Maricopa County, Arizona | Phoenix-Mesa-Scottsdale |
| 48201 | 2 | Harris County, Texas | Houston-The Woodlands-Sugar Land |
| 6037 | 3 | Los Angeles County, California | Los Angeles-Long Beach-Anaheim |
| 48113 | 4 | Dallas County, Texas | Dallas-Fort Worth-Arlington |
| 17031 | 5 | Cook County, Illinois | Chicago-Naperville-Elgin |
| 48439 | 6 | Tarrant County, Texas | Dallas-Fort Worth-Arlington |
| 53033 | 7 | King County, Washington | Seattle-Tacoma-Bellevue |
| 48453 | 8 | Travis County, Texas | Austin-Round Rock |
| 27053 | 9 | Hennepin County, Minnesota | Minneapolis-St. Paul-Bloomington |
| 48029 | 10 | Bexar County, Texas | San Antonio-New Braunfels |
| 13121 | 11 | Fulton County, Georgia | Atlanta-Sandy Springs-Roswell |
| 6073 | 12 | San Diego County, California | San Diego-Carlsbad |
| 32003 | 13 | Clark County, Nevada | Las Vegas-Henderson-Paradise |
| 48085 | 14 | Collin County, Texas | Dallas-Fort Worth-Arlington |
| 48121 | 15 | Denton County, Texas | Dallas-Fort Worth-Arlington |
| 8031 | 16 | Denver County, Colorado | Denver-Aurora-Lakewood |
| 37119 | 17 | Mecklenburg County, North Carolina | Charlotte-Concord-Gastonia |
| 12057 | 18 | Hillsborough County, Florida | Tampa-St. Petersburg-Clearwater |
| 39049 | 19 | Franklin County, Ohio | Columbus |
| 12095 | 20 | Orange County, Florida | Orlando-Kissimmee-Sanford |

*Notes*: A list of top ranked counties in the U.S. based on 2019 migration statistics using Geographic PageRank.