

## Discussion of "Measuring Housing Quality Using Revealed Preference"

By Jacob Krimmel

### **I. Introduction**

Measuring location quality is as important as it is difficult for the field of urban economics. It is a fundamental parameter in our models, yet because it is latent, multidimensional, and spatially correlated with many other variables, location quality is one of the most persistent confounders in our research, routinely producing wrong-signed hedonic estimates suggesting, for example, homebuyers pay a premium to live in more polluted or higher-crime areas.

What the field requires is a general-purpose measure of unobserved location quality that is transparent in its construction, grounded in economic behavior, and versatile enough to serve as an input across a range of identification strategies. In "Measuring Housing Quality Using Revealed Preference," Bell, Calder-Wang, and Zhong (2026) propose exactly this. The authors adapt the PageRank algorithm (Page 2001) to the U.S. household migration network, leveraging the intuition that households on average move from worse to better places and that locations drawing migrants from other highly-ranked origins are themselves revealed to be of high quality. The result is Geographic PageRank (GPR), a model-free ranking of latent place quality for U.S. counties and metropolitan areas.

As a proof of concept, the authors demonstrate that GPR can serve as an "anti-instrument" for unobserved quality in hedonic regressions, recovering a correctly signed housing price elasticity to air pollution of -0.40 that aligns with the quasi-experimental benchmark established by Chay and Greenstone (2005). This application alone illustrates the metric's potential to resolve long-standing identification failures in the literature.

This discussion commentary evaluates GPR not only as a descriptive index but as a general-purpose technology for urban economists. To structure that evaluation, I introduce a framework designed for the assessment of new economic measurements and apply it to this paper.

## II. A Framework for Evaluating Economic Measurement

Evaluating a measurement paper is a different exercise than evaluating a causal inference or model-based paper. A good metric is intuitive, replicable, and useful across a range of contexts. A good measurement *paper* therefore must communicate and prove that the new metric hits those marks. I propose a five-pillar framework for evaluating an economic measurement paper:

1. **Motivation:** Why do we need this measure? What is currently lacking in our data or toolkit?
2. **Innovation:** Is the measure actually novel? What is the conceptual or methodological leap over existing approaches?
3. **Explanation:** How is the measure constructed? Is the methodology transparent enough to replicate?
4. **Validation:** How do we know it works? Does it align with established benchmarks and possess predictive power?
5. **Application:** What are the use cases? How does this metric change the way we build models or identify causal effects?

The sequencing matters here. A measurement paper cannot validate what it has not explained, and validation must precede application for the same reason we calibrate models before running counterfactuals. The latter two pillars are also not solely the authors' responsibility — validation

and application are where the broader discipline must engage. My second and third sets of comments center on this.

Bell, Calder-Wang, and Zhong (2026) thoroughly satisfy Motivation and Explanation. The motivation is well-established: existing approaches to measuring location yield the wrong sign in hedonic regressions. The explanation is equally strong: the algorithm computes a stationary distribution of the migration transition matrix, weighting destinations more heavily when they draw migrants from other highly-ranked origins, and the methodology is transparent, publicly documented, and accompanied by a downloadable data product.

The remainder of this commentary focuses on the Innovation, Validation, and Application of GPR. This is where future research should push hardest. In particular, I highlight how economic frictions, like housing supply constraints which increase the cost of living, geographic distance which increases moving costs, and historical migration patterns which underly the data, all shape the migration flows on which GPR relies. Understanding the direction and magnitude of the biases these frictions introduce is essential for researchers who intend to use GPR as a structural input. The metric measures where people actually go, and that is not the same as where they would go in a frictionless world.

### **III. Innovation: Frictions, Constraints, and the Bias Structure of GPR**

The conceptual starting point is Sorkin (2018), who used PageRank to rank firms based on worker-to-worker transitions. The logic was simple: excluding layoffs, workers on average move from worse to better employers, so the network of transitions reveals firm quality. Bell, Calder-Wang, and Zhong (2026) apply the same recursive logic to households and locations.

However, the transition from a network of firms to a network of geographic locations introduces frictions that are qualitatively different, and quantitatively far larger, than those governing labor market transitions. In the labor market, workers actively search for jobs within local labor markets at relatively low cost; switching employers within a metropolitan area involves negligible physical relocation. In the housing market, by contrast, households are inherently endowed with a location, and the choice set across counties or metropolitan areas is far more constrained. The relevant frictions are massive. Between 2010 and 2019, the average annual quit rate in the U.S. labor market was approximately 3.8 percent, whereas the cross-county migration rate hovered around 2.4 percent. Though these are close in percentage point terms, the numbers also imply people are 60% more likely to quit their job – an already rare event – than move across county lines. Moving across metropolitan areas involves physical, financial, and social transaction costs, including the direct costs of relocation and the dissolution of local social networks.

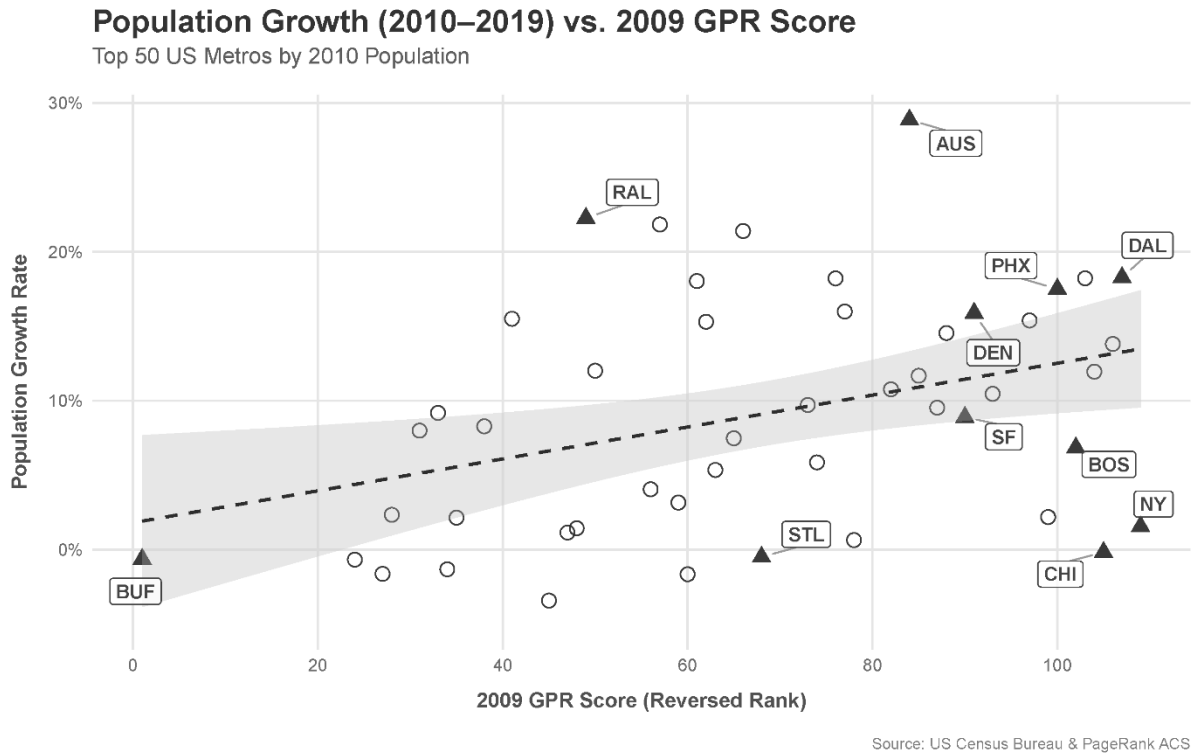
These frictions, which have no direct analogue in a within-metro job switch, matter because they filter who actually shows up in the migration data. GPR does not observe latent demand for a location. It observes only realized moves—the subset of potential migrants who clear the relevant friction. This means GPR is subject to identifiable biases, and supply constraints are the most consequential source of bias for researchers working in U.S. housing markets.

### *Supply Constraints as the Central Bias*

Perhaps the most significant friction shaping U.S. migration flows is the severe restriction on housing supply imposed by local residential land-use regulations. In highly productive coastal markets, strict zoning regimes and regulatory barriers severely limit the elasticity of housing supply, capitalizing excess demand directly into prices rather than permitting it to manifest as

population growth (Glaeser and Gyourko 2018; Hsieh and Moretti 2019). The empirical implications for GPR are immediate.

Figure 1: Scatter plot of 2021 GPR Score vs. Population Growth (2010–2019).



High-elasticity Sunbelt markets (e.g., Dallas, Phoenix, Houston) denoted with solid black triangles; severely supply-constrained coastal markets (e.g., San Francisco, New York, Boston) denoted with hollow circles.

Consider Figure 1, which plots 2021 GPR scores against the prior decade's population growth for top U.S. metros. Two empirical patterns are immediately visible. First, high-elasticity Sunbelt markets (e.g. Houston, Dallas, Phoenix, Austin) top the GPR rankings alongside massive population growth. These markets permit demand to be absorbed through new construction, and both the price signal and the migration signal move in the same direction. Second, and more

critically, severely supply-constrained coastal markets (e.g. San Francisco, Boston, New York) maintain high GPR scores despite stagnant or negative population growth. In these markets, the regulatory environment prevents demand from being expressed through new housing starts, so demand is instead capitalized into prices that suppress the volume of successful in-migration.

This pattern has two important implications for how researchers should interpret GPR. First, GPR is likely downward-biased for supply-constrained markets. In a frictionless world where San Francisco could freely expand its housing stock, migration inflows would be substantially larger, and its GPR rank would be correspondingly higher. The algorithm only observes the migrants who successfully clear the affordability hurdle; the marginal household that would have moved to San Francisco but was priced out is invisible to the metric. This is a selection-on-the-margin problem: GPR captures the intensive margin of realized moves but not the extensive margin of suppressed demand.

Second, the divergence between a market's GPR rank and its observed population growth may itself be informative about the severity of binding local supply constraints. Under the assumption that GPR captures a reasonably accurate signal of latent demand conditional on the frictions that filter migration, a market with high GPR but low population growth is one where demand substantially exceeds the housing supply response. The magnitude of this divergence, that is, the residual between what the network algorithm implies about a market's desirability and what the population data show about actual absorption, should be proportional to the degree to which local land-use regulations bind. Empirically, one could test this conjecture by correlating the GPR-growth residual with direct measures of regulatory stringency, such as the Wharton Residential Land Use Regulatory Index (Gyourko, Hartley, and Krimmel 2021). If the residual systematically correlates with measured regulatory burden, it suggests that GPR, in combination

with population data, can serve as a revealed-preference indicator of binding supply constraints – a useful diagnostic in its own right, independent of the metric's intended application.

### *Distance Penalties and the "Sydney Problem"*

Geographic distance works like a demand-side analogue to supply constraints. The migration literature documents a robust negative relationship between distance and the probability of moving (Kennan and Walker 2011). Consider what I call the "Sydney Problem": if Sydney, Australia, were a four-hour drive from the U.S. interior, migration flows would reveal it as a top-tier destination. Because it is thousands of miles away, its network ties are sparse, and GPR assigns it a low rank. Distance does to Sydney what zoning does to San Francisco—it suppresses the flows that would otherwise reveal true preference. Within the continental U.S., the effect is less extreme but still present: geographically remote markets may rank lower than their fundamentals warrant.

### *Inertia in Migration Networks*

A third source of bias arises from the persistence of historical migration networks. As Schubert (2024) demonstrates, migration pipelines between U.S. cities are remarkably stable over time: households tend to follow existing social and familial networks when making location choices, independent of the destination's current amenity value. This inertia is reflected in the GPR rankings, which exhibit high serial correlation: the correlation between 1990 and 2021 rankings sits at 0.88. If a substantial fraction of the migration signal driving GPR reflects the gravitational pull of pre-existing network ties rather than a contemporaneous assessment of location quality, then the metric may overstate the rank of historically popular destinations and understate the rank of emerging ones.

Network inertia creates a bias distinct from supply constraints and distance: rather than suppressing the level of migration, it biases the composition of flows by channeling households along established corridors regardless of whether those corridors still lead to the highest-quality destinations. The validation exercises I discuss in the next section are designed, in part, to disentangle these sources of bias. <sup>1</sup>

#### **IV. Validation: Predictive Power, Inertia, and the Limits of Ordinal Rank**

GPR passes the descriptive "eye test" It captures high demand along the coasts and rapid growth across the Sunbelt, and it correlates with net migration while retaining independent variation. But confirming what we already know (descriptively validity) is naturally in tension with the claim of innovation. If GPR reproduces patterns visible in simpler metrics, the algorithmic complexity has not earned its keep. I propose three validation exercises tied to the bias structure from Section III.

---

<sup>1</sup> The biases discussed above are primarily microeconomic frictions operating on individual migration decisions. But GPR also faces a macroeconomic measurement challenge: because the metric relies on gross dynamic migration flows, its precision varies with aggregate mobility. When macroeconomic transaction costs spike---as when mortgage rates exceeded 7 percent in 2022, freezing cross-county migration through the lock-in effect (Aladangady, Krimmel, and Scharlemann 2025)---the transition matrix is estimated from a thinner set of flows, potentially introducing noise or systematic bias toward locations that attract movers least sensitive to interest rate conditions. Understanding how these macro regimes influence the real-time precision of GPR will be important for researchers using the metric across different time periods.

### *Predictive Horseraces*

If GPR captures unobserved quality better than simpler alternatives, it should predict future housing market outcomes better too. Researchers should regress future house price appreciation—at a five-year horizon, say—on current GPR rank, net migration, and gross inflows, and test whether GPR adds marginal predictive power. A parallel exercise using future population growth would also be informative. GPR does not need to be a perfect predictor. It just needs to beat the alternatives that ignore the recursive network structure.

### *Disentangling Quality from Network Inertia*

A second validation exercise targets the network inertia bias directly. The high serial correlation in GPR rankings noted above could reflect genuinely durable fundamentals or sticky migration pipelines. To disentangle these channels, researchers should investigate how GPR responds to large, discrete shocks to local fundamentals. Natural disasters, major plant closings, or the sudden arrival of a large employer generate sharp, exogenous changes in a location's desirability. If GPR adjusts promptly to such shocks, the serial correlation likely reflects genuine quality persistence. If the rankings are sluggish---if a location that suffers a major negative shock retains a high GPR rank for several years due to legacy migration flows---that would suggest the metric needs to be interpreted with a longer-run lens, and that short-run changes in rank may be noisy signals of quality shifts.

### *The Limitations of Ordinal Rank*

Finally, researchers must grapple with the natural limitations of ordinal rankings. Consider the "Fallers" in the data: Denver drops 24 rank positions between 2001 and 2021, while Sacramento drops 14. Did Denver's absolute quality actually decline that much, or did other close substitute

markets leapfrog it by building more housing? If high-elasticity Sunbelt metros absorb population through new construction, they can jump past supply-constrained markets in the rankings without the constrained market getting any worse. This is a direct consequence of the supply-constraint bias from Section III: rank changes conflate genuine quality shifts with differential supply dynamics. One productive path forward would decompose CBSA rank changes over time into a component explained by local permitting activity and population absorption (the supply channel) and a residual component that plausibly reflects changes in amenities or labor market conditions (the quality channel). Without such decomposition, the ordinal nature of GPR introduces an interpretive ambiguity that will limit its usefulness in applications requiring cardinal measures of quality.

## **V. Application: Spatial Equilibrium, Granularity, and Market Segmentation**

The risk with any new index is replacing one black box with another. I argue GPR's value lies not in the rankings themselves but in what they enable as an input to structural models and identification strategies.

### *GPR as an Input to Spatial Equilibrium*

Beyond the authors' anti-IV application, the most consequential use of GPR will be its integration into the Rosen-Roback spatial equilibrium framework (Rosen 1979; Roback 1982). GPR captures a composite of housing prices, labor markets, and amenities in a model-free way, making it a natural input for models that need a measure of unobserved location quality.

However, this creates an inherent economic tension. High amenity does not equal high quality of life if housing costs consume the entirety of the local wage premium (i.e. housing is chronically unaffordable). A market that ranks highly on GPR because it offers exceptional amenities may

simultaneously deliver low welfare to its residents if those amenities are fully capitalized into rents. Future researchers must use GPR alongside localized wage indices and rent-to-income ratios to decompose how much of a market's rank is driven by pure exogenous amenity value (think San Francisco) versus endogenous affordability (think Dallas). The supply-constraint framework developed in Section III is directly relevant here: in constrained markets, a disproportionate share of the amenity value is capitalized into prices rather than absorbed through population growth, and the GPR rank reflects this constrained equilibrium rather than a welfare-relevant quality measure.

#### *How Granular Can GPR Go?*

The authors note that the methodology can be extended to finer scales with restricted census data. This is an important frontier, but the challenges are not just about data availability.

As the geographic unit shrinks toward the census tract, the migration network becomes sparse and the transition matrix is populated with zeros. The stationary distribution may be poorly identified. More fundamentally, the revealed-preference logic weakens: a move from one tract to an adjacent one may reflect a change in unit size or the availability of a listing, not a deliberate preference over neighborhood quality. Spatial spillovers also complicate things: a tract's desirability depends on the quality of its neighbors through schools, transit, and local externalities. Whether PageRank can capture these dependencies at fine scales is an open question. If these challenges can be addressed, tract-level GPR would let researchers price hyper-local amenities with far greater precision than metro-level rankings allow.

#### *Housing Market Segmentation by Population Subgroup*

One of the paper's most promising extensions is the computation of GPR by population subgroup using American Community Survey data. The authors demonstrate that subgroup rankings are generally positively correlated but that important differences emerge---particularly by race and ethnicity, where correlations drop as low as 0.74. These divergences are substantively meaningful and open a productive line of research on how the spatial equilibrium differs across demographic groups.

Computing separate rankings by tenure status (owners versus renters) is a particularly informative extension. Renters face substantially lower moving frictions than homeowners: they are not anchored by existing home equity, do not forfeit a favorable mortgage rate by relocating, and face lower transaction costs. A renter-specific GPR may therefore reveal latent demand more cleanly than a GPR computed from all movers, because the migration signal is less attenuated by the financial frictions that filter owner mobility. Comparing owner and renter GPR rankings across markets can help quantify how lock-in effects and equity constraints segment the housing market, and where latent demand for locations is most suppressed by the frictions of homeownership.

## **VI. Conclusion**

Bell, Calder-Wang, and Zhong (2026) have provided a transparent, rigorous, and versatile addition to the spatial econometrics toolkit. Geographic PageRank moves the literature forward from arguing over *how* to measure unobserved place quality to exploring *what* we can do with it.

The five-pillar framework applied here (Motivation, Innovation, Explanation, Validation, Application) is designed to be useful beyond this paper. As the discipline increasingly relies on

novel data products to resolve long-standing identification problems, we need a shared language for evaluating whether new metrics are ready for use as structural inputs.

As for the GPR metric, the most important interpretive point is also the simplest: GPR captures revealed preference *subject to constraints*. Migration networks are equilibrium outcomes, not first-best demand. Provided researchers account for the structural constraints that shape the underlying data, GPR will serve as a foundational metric for the next generation of urban economic research.

### **References:**

Aladangady, Aditya, Jacob Krimmel, and Tess Scharlemann. 2025. "Locked In: Mobility, Market Tightness, and House Prices." Finance and Economics Discussion Series 2024-088r1.

Washington: Board of Governors of the Federal Reserve System.

Bell, Alex, Sophie Calder-Wang, and Shusheng Zhong. 2026. "Measuring Housing Quality Using Revealed Preference." NBER Working Paper.

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

Glaeser, Edward, and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives* 32 (1): 3–30.

Gyourko, Joseph, Jonathan Hartley, and Jacob Krimmel. 2021. "The Local Residential Land Use Regulatory Environment across U.S. Housing Markets: Evidence from a New Wharton Index." *Journal of Urban Economics* 124: 103337.

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11 (2): 1–39.

Kennan, John, and James R. Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica* 79 (1): 211–251.

Page, Lawrence. 2001. Method for Node Ranking in a Linked Database. U.S. Patent 6,285,999, filed January 9, 1998, and issued September 4, 2001.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–1278.

Rosen, Sherwin. 1979. "Wage-Based Indexes of Urban Quality of Life." In *Current Issues in Urban Economics*, edited by Peter Mieszkowski and Mahlon Straszheim, 74–104. Baltimore: Johns Hopkins University Press.

Schubert, Gregor. 2024. "House Price Contagion and U.S. City Migration Networks." Working Paper, UCLA Anderson School of Management.

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