Cheaper but Farther: Rent Gradients and Commuting in the Residential Choices of Young Korean Workers

Hyun Hak Kim*

September 2026

Abstract

Where do young wage workers live, given where they work? This paper studies the residence municipality chosen by employed wage workers aged 19–34 in Korea, conditional on their observed workplace, using survey microdata for 2021H2–2024H2 linked to a standardized rent index built from administrative lease records for every rent-eligible municipality. In a conditional logit over all alternatives, the log-rent coefficient is −1.153, which implies a −10.4 percent change in the relative odds of a residence with 10 percent higher rent; it is stable across choice sets, survey rounds, and inference methods. A within-between decomposition locates the gradient in persistent differences between municipalities (−1.254) rather than in changes within them (0.080). The between-municipality gradient is confined to capital-region residences (contrast −2.166, and −1.903 after the noncapital cities that dominate cross-boundary commuting are removed), and workers at high-rent workplace centers combine a steeper gradient with cheaper, more distant residences. The estimates are conditional associations, not responses to an exogenous change in rent; they show that persistent rent differences order the residential choices of young workers attached to capital-region jobs, where housing, transport, and job policies need to be designed together.

Keywords: youth residential choice; housing rents; spatial sorting; commuting; capital-region concentration; Korea
JEL codes: J61, R21, R23, R41

*Corresponding author. Professor, Department of Economics, Kookmin University, Seoul, Korea.

1 Introduction

Where young workers live is a joint outcome of housing costs, access to jobs, amenities, and housing supply, and no single coefficient in a residential-choice model separates these forces. Rent is an equilibrium price that capitalizes access and amenities (Roback, 1982; Glaeser and Gottlieb, 2009), so a negative coefficient on the rent of a residence may reflect cost avoidance, sorting on unmeasured quality, or the location of jobs, and housing costs in expensive labor markets may themselves limit who gets to work there (Hsieh and Moretti, 2019; Ganong and Shoag, 2017). Young workers are the group for whom these forces bind most. This paper does not try to separate all of these forces. It asks a narrower question that the available data can answer: among employed wage workers aged 19–34 who hold a job in a given municipality, how is the standardized rent of a municipality related to whether they choose it as a residence, once the workplace, commuting distance, and measured amenities are held fixed, and where in the country and for whom does that relation hold?

Korea makes the question sharp. The capital region, which comprises Seoul, Incheon, and Gyeonggi, holds 57.8 percent of the workplaces in the estimation sample of young wage workers, and its rent levels sit above the rest of the country in a stable order. Measured in 10,000 won per month at 2024 prices, the median standardized rent of a 40 m^2 unit across municipality-half-year cells is 63.8 in Seoul, 45.9 in Gyeonggi, and 42.4 in Incheon, against 35.5 outside the capital region (Figure 3). Jobs and high rents are therefore jointly concentrated, and a young worker attached to a capital-region job chooses a residence among municipalities whose rent differences are large and persistent. Whether the rent gradient is a national regularity or a feature of the capital region also decides whether housing costs for young workers are a national question or a metropolitan one. Throughout the paper the spatial unit is the municipality (*si-gun-gu*), the time unit is the half-year of the survey, and the outcome is the residence municipality chosen by wage workers with an observed workplace.

The design conditions on the observed workplace. For each worker the alternatives are every municipality in the country whose lease records support a measured rent in that half-year, 207–212 municipalities, and the choice among them is modeled as a conditional logit (McFadden, 1974). Because the workplace is fixed, its wages, job openings, and industry mix are common to all alternatives and drop out of the comparison; what remains is a comparison of residences by rent, amenities, and distance to that workplace. The full choice set makes sampling of alternatives unnecessary, and sampled choice sets with the appropriate correction serve only as a check (McFadden, 1978). The rent measure is a standardized index built from administrative lease contracts, so rent differences across municipalities are measured on a common basis rather than from the workers' own housing costs. Separating the persistent level of a municipality's rent from its half-year changes matters in this setting, because level differences are what a worker attached to a capital-region job faces when choosing among municipalities, whereas half-year changes mix local shocks with the composition of reported leases.

Three questions organize the analysis. Is the conditional rent gradient negative, and does it survive changes in the choice set, the survey round, and the inference method? Does the gradient come from persistent differences in rent levels between municipalities or from changes in a municipality's own rent over time? And is the gradient a national regularity or a feature of the capital region, and how does it relate to the commuting of workers attached to expensive workplace centers? Differences in the gradient by age, sex, and education were predeclared as uncertain in sign and are reported as an auxiliary test.

The data link seven half-years of the Regional Employment Survey of Statistics Korea, 2021H2–2024H2, to a rent index built from the lease records of the Real Transaction Management System (RTMS). The estimation sample contains 224,862 employed wage workers aged 19–34 with residence and workplace municipality codes and a rent-eligible residence (Table 1). A hedonic model estimated on 5,159,595 monthly-rent contracts from 2017H1 to 2024H2 gives the rent of a standard 40 m^2 unit for each municipality and half-year, and the rent-eligible cells of the analysis window contain 3,102,349 contracts. The empirical strategy proceeds in steps: a Mundlak within-between decomposition of rent and municipality fixed effects with trends, a residence-by-workplace market model in which rent slopes differ between capital-region and noncapital residences and workplaces, commuting tests around high-rent workplace centers, and a reweighting exercise for employment selection on observed characteristics. Inference uses two-way clustering by residence and workplace geography and wild-score intervals.

Three results follow. The conditional log-rent coefficient in the full choice set is −1.153 (0.263), which implies a −10.4 percent change in the relative odds of a residence whose standardized rent is 10 percent higher; the estimate does not depend on sampled alternatives, random seeds, any single half-year, or the inference method. The gradient arises from persistent between-municipality differences: the within-municipality coefficient is 0.080 ($p = 0.273$), whereas the between-municipality coefficient is −1.254 ($p < 0.001$). And the between-municipality gradient is confined to capital-region residences. The capital-minus-noncapital contrast is −2.166 ($p < 0.001$) for all workers and −1.903 after the five largest noncapital cities near the capital boundary are removed from every choice set, and no gradient is detected across noncapital residences. Among workers employed in the capital region the contrast is larger, −5.531 (1.206), but that figure depends on those few cities and is reported only together with the exclusions. Consistent with sorting around expensive workplaces, workers at high-rent workplace centers have a rent slope steeper by -1.501 (0.357) than other workers and more often live in a municipality that is cheaper than their workplace but farther from it. Reweighting wage workers toward all persons aged 19–34 on observed characteristics does not overturn these gradients (Section 6).

The paper makes four contributions. It measures rent nationwide from administrative lease records at the municipality-half-year level and links the measure to survey microdata that record both residence and workplace. It conditions on the observed workplace and uses the full set of residence alternatives, so that the rent gradient is a residential comparison rather than a mixture of job and housing choices. It separates persistent rent levels from changes in rent within a municipality. And it reports the capital-region market structure together with a diagnosis of the few adjacent cities on which its sharpest contrast rests. Relative to the Korean literature on youth location, which studies movement between broad regions or between college and workplace, the object here is the residence of an employed worker conditional on the workplace, chosen from every rent-eligible municipality in the country. The estimates are conditional associations among employed wage workers; the paper does not identify a causal response to rent, does not measure workers' own housing costs, and does not evaluate welfare. The implication it does support is a design principle: youth housing, transport, and job policies should be considered together with the structure of capital-region workplace markets. Section 2 places the design in the literature and states the predeclared predictions, Section 3 describes the data, Section 4 sets out the empirical strategy, Section 5 reports the main results, Section 6 reports the robustness checks and the employment-selection sensitivity, and Section 7 concludes.

2 Conceptual Framework and Related Literature

This section places the design in four strands of literature and closes with the predictions fixed before estimation. The strands are treated together rather than as separate theory and literature sections, because each supplies one element of the design: what a rent gradient measures, why housing costs matter for access to productive labor markets, how commuting separates residence from workplace, and how early-career workers in a single-metropolis economy choose where to live.

2.1 Spatial equilibrium and rent as a price of access

In the spatial-equilibrium tradition of Roback (1982), wages and land rents adjust until workers are indifferent across locations, so local rent capitalizes the amenities and the productive opportunities that a location offers (Glaeser and Gottlieb, 2009; Moretti, 2011). Two consequences follow for a residential-choice model. First, the coefficient on residence rent is not the price of a homogeneous good. Holding the workplace fixed removes the wage side of the equilibrium for a given worker, but the residence side still bundles housing cost with amenities, access to other jobs, and unobserved neighborhood quality that is capitalized into rent. A negative rent coefficient therefore measures the net of cost avoidance and capitalization, and a small coefficient does not show that housing costs are irrelevant. Second, the equilibrium logic distinguishes persistent from transitory variation in rent. Long-run differences in rent levels across municipalities reflect the durable structure of amenities and access, whereas half-year deviations from a municipality's own level reflect local shocks, reporting composition, and short-run market conditions. The two sources of variation need not carry the same information about residential choice, which is why the empirical strategy separates within-municipality changes from between-municipality levels (Section 4.3). Heterogeneity by skill adds a further layer: Diamond (2016) shows that high-wage, high-rent cities can also offer endogenous amenities that attract particular groups, so sorting on rent and sorting on amenities are entangled.

2.2 Housing costs, mobility, and access to productive labor markets

A second strand asks whether housing costs limit access to productive places. Hsieh and Moretti (2019) argue that housing supply constraints in high-productivity cities lower aggregate output by keeping workers away, and Ganong and Shoag (2017) document that rising housing costs in high-income areas have weakened the historical link between income differences and directed mobility, particularly for less-skilled workers. Saks (2008) shows that where housing is hard to build, labor demand shocks translate into higher wages and less employment growth, so housing supply shapes how local labor markets absorb workers, and Moretti (2013) shows that differences in local prices change real-wage comparisons across skill groups. These studies concern the allocation of people across labor markets, and their identification rests on shocks to demand or supply. The present paper occupies a narrower position. It does not identify the causal response of location to housing costs, and it does not observe moves: the survey is a sequence of repeated cross-sections, so what is observed is the distribution of residences chosen by workers who hold a given job, half-year by half-year. This strand contributes the question, whether young workers attached to expensive labor markets are ordered across residences by persistent rent differences, and the caution that rent in such markets is an equilibrium outcome rather than a treatment.

2.3 Commuting and the separation of residence and workplace

The third strand treats commuting as the margin along which residence and workplace are separated. Quantitative urban models generate commuting flows from a gravity relation in which workers trade off wages at the workplace, housing costs at the residence, and the cost of travel between them (Ahlfeldt et al., 2015). Heblich et al. (2020) show that the steam railway produced the first large-scale separation of workplace and residence in London, and Monte et al. (2018) show that commuting links residence and workplace markets, so that local employment adjusts differently where commuting is easy. Manning and Petrongolo (2017) find that the attractiveness of a job falls steeply with distance, which makes the residence choice of a worker with a given job a spatially bounded problem. This strand motivates the conditioning structure of the empirical model. Because the workplace is fixed, the residence choice is a trade-off among municipalities that differ in rent, amenities, and distance to that workplace; the distance terms and the indicators for a residence in the same municipality or province as the workplace absorb the commuting cost, and the rent terms are read conditional on them. It also motivates the descriptive question of whether workers at expensive workplace centers combine cheaper residences with longer commutes. The limits of the strand carry over: distance is measured between municipality centroids, a coarse proxy for travel time, especially within large cities.

2.4 Early-career location choice and the Korean setting

The fourth strand concerns early-career workers. In the human-capital view of Sjaastad (1962), a location choice balances expected returns against the monetary and nonmonetary costs of moving, and Kennan and Walker (2011) show that expected income differences are a major component of location choices among young workers, alongside moving costs large enough to keep most people in place. Housing costs enter this calculus on the return side, as a deduction from real income at the destination, and on the cost side, as a barrier to entry. Young workers are the group for whom both channels are most active: they have the longest horizon over which returns accrue, the fewest housing assets, and the highest rates of job change.

Korea gives this strand a particular form. Jobs, high rents, and young workers are concentrated in a single metropolitan region around Seoul, so the distinction between capital-region and noncapital residence alternatives, and between capital-region and noncapital workplaces, is the natural place to look for market structure in the rent gradient. The Korean literature on youth location has mainly studied movement between regions rather than residence conditional on workplace. Shim and Kim (2012) analyze the employment mobility of college graduates between the capital region and the rest of the country and relate it to pay and job stability; Han and Yoon (2021) model the sequential choice of college and workplace regions and find that prior residence, distance, and local labor-market conditions matter at each stage; Choi (2022) links the social and cultural conditions of a region to graduates' departure from the region of their college; and Choi and Lee (2017) describe the youth labor market of a noncapital region and the inflow and outflow of graduates in terms of wage gaps and prior local ties. These studies take movement between broad regions, or between college and workplace, as the outcome. The present paper differs in the object of choice and in the conditioning. The outcome is the residence municipality of an employed wage worker with an observed workplace, chosen from every rent-eligible municipality in the country, and the variable of interest is a standardized rent index built from administrative lease records rather than a regional average. The difference is not that residence conditional on workplace has never been studied, but that the combination of nationwide rent measurement, workplace conditioning, the within-between decomposition, and the capital-region market structure has not, and that combination is what allows the paper to say where, and for whom, persistent rent differences order residential choice.

2.5 Predeclared predictions

Four predictions were fixed before estimation, and the results are read against them rather than against a list of hypotheses. First, conditional on the workplace, the between-municipality rent gradient is negative: among the municipalities that a worker with a given job could choose, those with persistently higher standardized rent are chosen less often, other measured attributes held fixed. Second, the within-municipality rent gradient may be smaller in absolute value than the between-municipality gradient, because half-year deviations in a municipality's rent are less likely to reorder choices than durable level differences. Third, the rent gradient is more negative for capital-region residence alternatives than for noncapital alternatives, and the difference is larger among workers employed in the capital region, where jobs and high rents are jointly concentrated. Fourth, differences in the gradient by age, sex, and education are uncertain in sign, and they are reported whether or not they are significant. The predictions concern conditional associations among employed wage workers with a given workplace. None of them is a claim that higher rent makes a worker avoid a location, and none extends to young people who are not employed.

3 Data

3.1 Sources and study period

The analysis links five sources at the level of the municipality and half-year. The Regional Employment Survey of Statistics Korea, accessed through its Microdata Integrated Service (MDIS), records the residence and workplace municipality of each respondent. Lease contracts come from the Real Transaction Management System (RTMS) of the Ministry of Land, Infrastructure and Transport. Youth job openings come from the Work24 administrative statistics of the Ministry of Employment and Labor. Local service indicators come from the e-Local Indicators of the Korean Statistical Information Service (KOSIS), and municipality centroids and boundaries come from the Statistical Geographic Information Service (SGIS).

The spatial unit is the municipality (*si-gun-gu*). The survey's semiannual round, called a half-year (*semester*) here, distinguishes 229 municipalities in every round, each observed both as a residence and as a workplace. A bridge links survey codes, RTMS legal-district codes, and SGIS boundary units to reconcile administrative changes, and a stable-geography identifier is kept for the decomposition and for clustering. Appendix Table A1 lists the bridge rules. Figure 1 maps the provinces, the capital region, the commuter cities of Section 5.3, and the rent level of every municipality.

The study period is 2021H2–2024H2, which gives seven repeated cross-sections. Two considerations fix this window. Residence and workplace municipality codes are consistently available in the survey files for these half-years, and the period follows the start of mandatory lease reporting in June 2021, which made RTMS monthly-rent records usable for a nationwide measure of small-unit rent. Because each half-year is a new cross-section, the analysis describes where employed young workers live in each half-year rather than how individuals move.

3.2 Worker sample and observed choices

Table 1 traces the construction of the sample. The survey records 257,262 employed persons aged 15–34 over the seven half-years. The choice analysis keeps wage workers aged 19–34 with valid survey weights, sex, education, residence, and workplace, which leaves 229,925 observations. The lower age bound excludes minors and aligns the sample with the end of secondary schooling. A residence must also have a rent measure (Section 3.3), and this requirement removes 5,063 workers. The estimation sample therefore contains 224,862 employed wage workers aged 19–34. Their survey weights sum to 39,650,603 over the seven cross-sections, a total that should not be read as a count of distinct people.

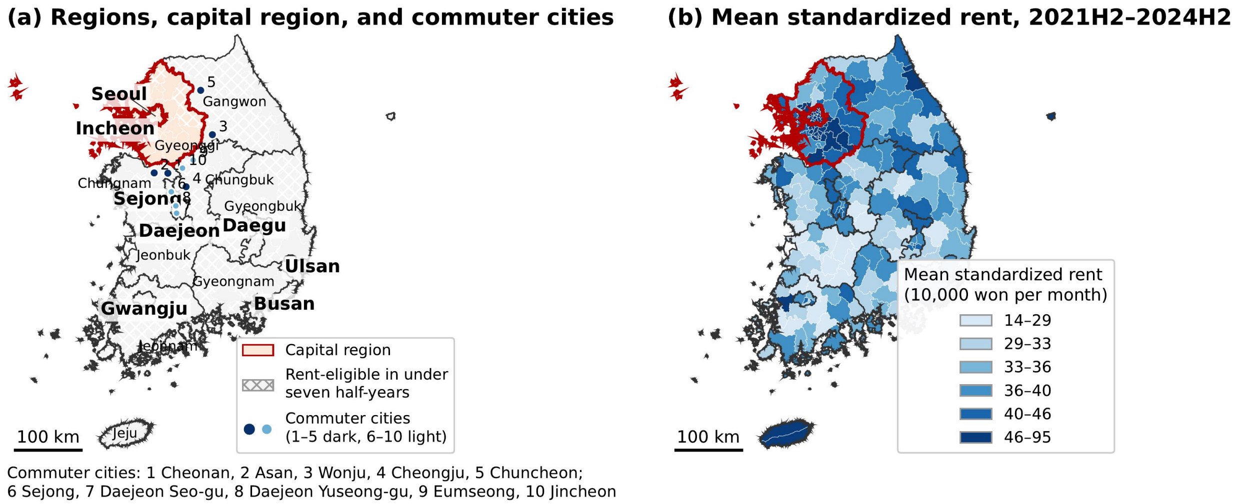

Figure 1: Regions, commuter cities, and standardized rent across Korea

Notes: Panel (a) shows the provinces and metropolitan cities with the capital region (Seoul, Incheon, and Gyeonggi) outlined in red. Hatched municipalities are rent-eligible in fewer than seven of the half-years 2021H2–2024H2. Numbered points mark the ten noncapital municipalities that carry the largest survey weight among noncapital residents employed in the capital region (Section 5.3); the first five are the commuter cities excluded in the robustness checks. Panel (b) shades each municipality by the mean of its standardized 40 m^2 monthly-equivalent rent over its rent-eligible half-years, in 10,000 won per month at 2024 prices; the class breaks are sextiles of the municipality means. Boundaries are the 2025 Q2 files of the Statistical Geographic Information Service (SGIS); a municipality that changed codes during the period is drawn with its latest boundary.

Table 1: Data sources, sample construction, and coverage

Panel A. Sources

	Unit and period	Used for
MDIS Regional Employment Survey (A)	Person, 2021H2–2024H2	Residence, workplace, worker type, weights
MOLIT RTMS lease contracts	Contract, municipality by half-year	Standardized 40 m^2 monthly-equivalent rent
Work24 administrative statistics	Municipality by half-year	Youth job openings (mechanism tests)
KOSIS e-Local Indicators	Municipality by year	Medical, cultural, education, transport indices
SGIS boundaries and centroids	Municipality	Distances and boundary bridges

Panel B. Sample flow

	Persons	Survey-weighted total
Employed persons aged 15–34	257,262	43,996,172
Wage workers aged 19–34	229,925	39,822,421
Less: residence rent cell below 20 contracts	5,063	171,818
Estimation sample	224,862	39,650,603

Table 1: Data sources, sample construction, and coverage (continued)

Panel C. Coverage by half-year

	Rent-eligible municipalities	Contracts in eligible cells	Choice groups	Positive flow cells	Persons
2021H2	211	394,658	3,107	11,331	31,392
2022H1	211	462,001	3,057	11,184	30,715
2022H2	212	425,844	3,124	11,750	34,327
2023H1	207	455,015	3,036	11,460	33,427
2023H2	209	441,348	3,055	11,357	32,775
2024H1	212	477,935	3,012	11,097	31,369
2024H2	211	445,548	2,988	10,910	30,857
All	n.a.	3,102,349	21,379	79,089	224,862

Notes: Panel B follows the Regional Employment Survey records from employed persons aged 15–34 to the wage-worker estimation sample. Survey-weighted totals sum over seven half-year cross-sections and are not counts of distinct people. A residence municipality enters the choice set in a half-year only if it has at least 20 eligible monthly-rent contracts. Panel C counts type-by-half-year-by-workplace choice groups and residence-workplace cells with positive observed flows in the full-choice data.

For estimation, the weights are rescaled to sum to the sample size. The rescaling keeps relative weights and prevents the population total of repeated cross-sections from acting as the number of independent observations. Inference relies on geographic clustering (Section 4.6).

Workers are grouped into types by sex, education, and age band (19, 20–24, 25–29, and 30–34). A choice group is a worker type employed in a given workplace in a given half-year. The full-choice data contain 21,379 choice groups. Summing survey weights over residence-workplace pairs gives 79,089 cells with positive observed flows. Because every group faces all rent-eligible residences, the full-choice data have 4,498,851 rows. Residence attributes in the core models do not vary by worker type, so the flows can be summed within half-year, workplace, and residence without changing the likelihood. This exact collapse yields 335,839 rows in 1,596 half-year-by-workplace groups.

Panel A of Table 2 sets the estimation sample beside all employed persons aged 15–34, who serve as context. In the estimation sample, 51.8 percent of workers live and work in the same municipality, and 85.2 percent live and work in the same province. The capital region holds 57.8 percent of workplaces in the estimation sample. The mean centroid distance between residence and workplace in the estimation sample is 8.6 km, and its 90th percentile is 24.5 km. The context sample of employed persons aged 15–34 is similarly attached, with 54.1 percent living and working in the same municipality. Figure 2 shows the same comparison together with the distribution of commuting distances and the largest cross-province flows.

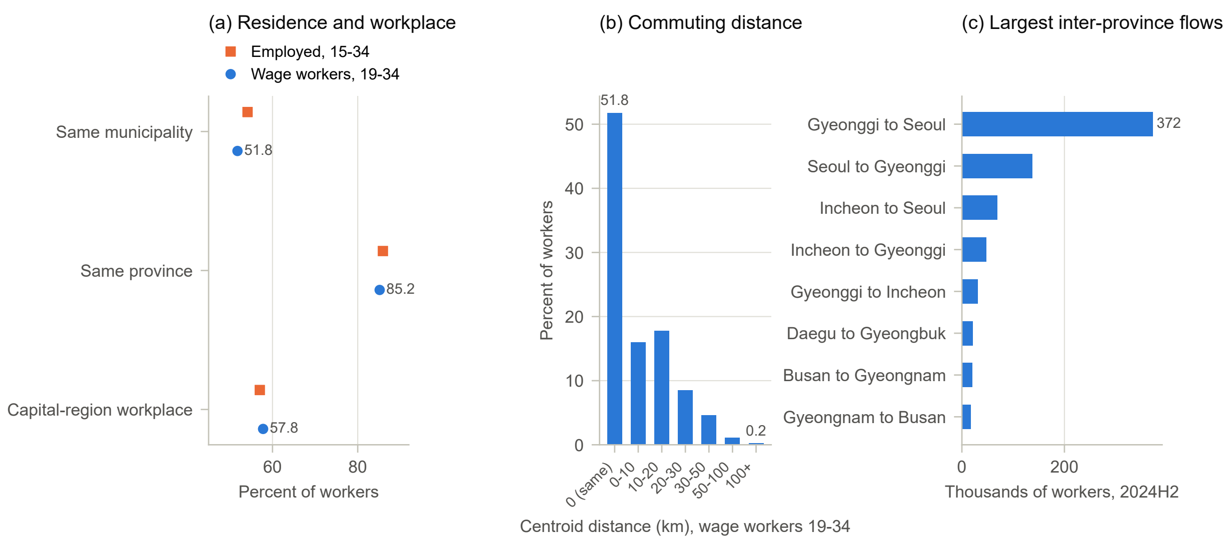

Figure 2: National residence-workplace geography

Notes: Panel (a) compares all employed persons aged 15–34 with the wage-worker estimation sample; dots pool seven half-years and lines span the half-year range. Panel (b) uses survey-weighted estimation-sample flows; distance is zero within a municipality. Panel (c) shows the eight largest cross-province flows of employed persons aged 15–34 in 2024H2.

Table 2: Residence-workplace geography and rent support

Panel A. Spatial structure

	Employed 15–34 (context)	Wage workers 19–34 (estimation)
Live and work in same municipality (%)	54.1	51.8
Live and work in same province (%)	85.9	85.2
Work in capital region (%)	57.0	57.8
Live in capital region (%)	n.a.	57.9
Mean commuting distance (km)	8.2	8.6
90th percentile distance (km)	n.a.	24.5
Mean distance, cross-municipality (km)	n.a.	17.8

Panel B. Standardized rent

	Cells	P5	Median	P95	Chosen median	Chosen geometric mean
Seoul	175	50.2	63.8	87.2	63.2	63.8
Incheon	70	34.0	42.4	56.9	44.2	44.9
Gyeonggi	217	32.8	45.9	64.3	48.0	46.8
Noncapital	1,011	23.0	35.5	44.6	37.7	37.0

Notes: Panel A pools seven half-years with survey weights. The context column uses all employed persons aged 15–34; the estimation column uses wage workers aged 19–34 in rent-eligible residences. Distances are between municipality centroids and are zero within a municipality; percentiles of the context sample cannot be pooled from half-year summaries. Panel B reports standardized 40 m^2 monthly-equivalent rent in 10,000 KRW at 2024 prices. Cell percentiles weight each rent-eligible municipality-half-year equally; chosen-residence statistics weight workers by survey weights. The geometric mean is the exponential of the mean log rent.

3.3 Standardized monthly-equivalent rent

The rent measure uses RTMS monthly-rent contracts for units no larger than 60 square meters. Pure deposit leases (*jeonse*) carry no monthly rent and are excluded, which targets the segment most relevant to young renters. A deposit-inclusive variant that admits *jeonse* leases, converting their deposits at the same market rate, serves as a sensitivity measure (Section 6.1; Appendix Table A1, Panel F). For contract h signed in month m, the monthly-equivalent cost is

$$MEC_{hm} = R_{hm} + \frac{i_m}{12} D_{hm},$$

where R_{hm} is the monthly rent, D_{hm} is the deposit, and i_m is the three-year Korean Treasury yield from the Bank of Korea. Costs are expressed in 2024 prices with the national consumer price index. Fixed conversion rates and monthly rent alone serve as alternative measures.

Raw medians would mix local price levels with differences in unit size, housing type, building age, floor, and contract month. The index therefore comes from a contract-level hedonic regression,

$$\log MEC_{hqs} = \alpha_{qs} + f(A_h) + g(B_h) + k(F_h) + \boldsymbol{\pi}'\mathbf{T}_h + \boldsymbol{\omega}'\mathbf{M}_h + u_{hqs},$$

where q is an RTMS geographic unit, s is a half-year, A_h is floor area, B_h is building age, F_h is the floor, $\mathbf{T}_h$ and $\mathbf{M}_h$ contain housing-type and month-within-half-year indicators with coefficient vectors $\boldsymbol{\pi}$ and $\boldsymbol{\omega}$, and u_{hqs} is an error term. The functions f, g, and k are quadratic, with indicators for missing values, and the outcome is winsorized within housing type at the 0.1st and 99.9th percentiles. The regression uses all 5,159,595 eligible contracts signed from 2017H1 to 2024H2, so the coefficients on unit characteristics are estimated on the full contract series. The fixed effect α_{qs} captures the price level of geographic unit q in half-year s. The standardized rent evaluates the fitted price at 40 square meters, ten years of building age, and the third floor, weighting housing types by fixed national contract shares. Holding the unit fixed prevents shifts in the composition of reported contracts from moving the index.

When several RTMS units map to one survey municipality, their standardized log rents are averaged with fixed weights proportional to each unit's contracts over 2021H2–2024H2. Fixed weights prevent changes in reporting volume across component districts from creating artificial movements in municipality rent. A municipality enters the residence choice set in a half-year only if it records at least 20 eligible contracts in that half-year, and no regional average is imputed for sparse cells. The rent-eligible cells of the analysis window contain 3,102,349 contracts. Each half-year's choice set contains 207–212 municipalities.

The threshold mostly removes rural counties, where lease reporting is incomplete. The reporting obligation introduced in June 2021 does not apply to counties (*gun*) in provinces or to small leases.[1] Of the municipality-half-year cells below the threshold, 128 lie in counties outside the mandatory reporting area and 2 lie inside it. Appendix Table A2 (Panel D) re-estimates the main models with thresholds of 10 and 5 contracts.

Across comparable cells, the standardized market-rate index has a correlation of 0.849 with the raw median monthly-equivalent rent. Appendix Table A1 reports the hedonic coefficients, the housing-type models, and further comparisons, including one with the deposit-adjusted rent series of the Korea Real Estate Board (R-ONE). These comparisons support the index as a ranking of local rent levels, but they do not turn an equilibrium rent into an exogenous price. Figure 3 shows standardized rent by region over time and the distribution of contract support across cells. Appendix Figure A2 maps the index within the capital region.

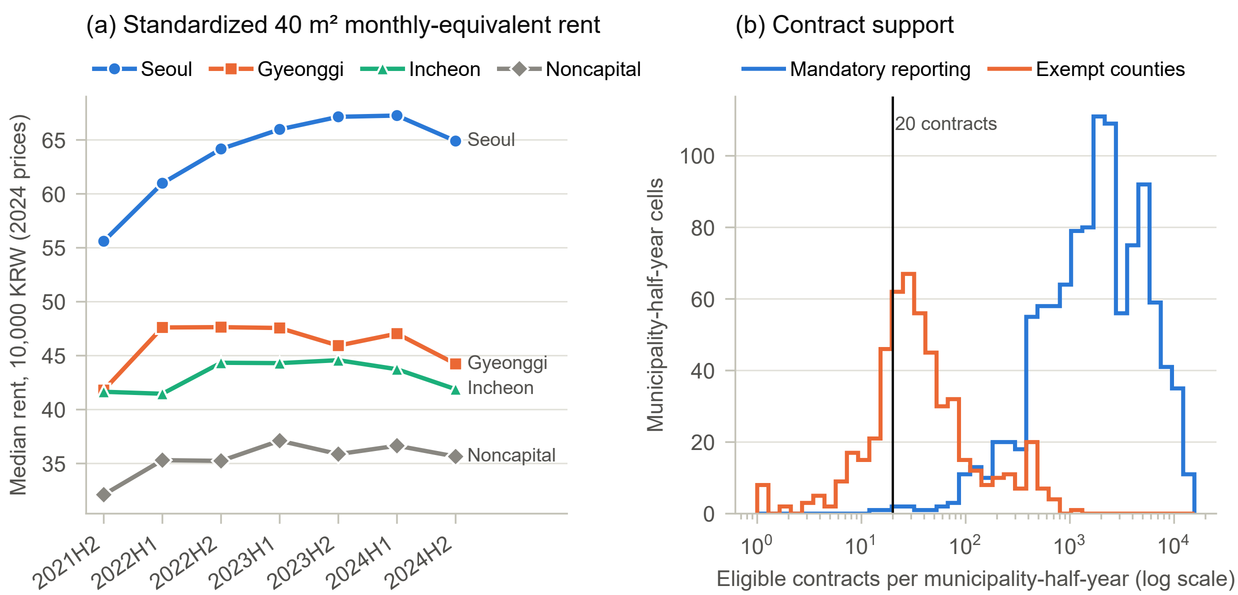

Figure 3: Construction and support of standardized rent

Notes: Panel (a) plots the median across rent-eligible municipalities (at least 20 contracts) of hedonic 40 m^2 monthly-equivalent rent at a market conversion rate and 2024 prices. Panel (b) counts municipality-half-year cells by eligible contracts; mandatory lease reporting since June 2021 excludes counties in provinces, and cells with no contract are shown at one.

3.4 Amenities, distance, and workplace measures

Four standardized indices from KOSIS describe local services. The medical index combines hospital beds and physicians per 1,000 residents, and the cultural index measures cultural

[1] Under the Act on Report on Real Estate Transactions, a residential lease must be reported when its deposit exceeds 60 million won or its monthly rent exceeds 300,000 won. Fines were not imposed during a guidance period that ran through May 2025.

facilities per 100,000 residents. The education index combines the number of universities with private academies per 1,000 residents, and the transport index is the road pavement rate. Where the survey municipality is a larger city, general districts are aggregated and ratios are recomputed from their components. The indices capture observed services, not local quality as a whole.

Commuting distance d_{rj} between residence r and workplace j is the great-circle distance between SGIS centroids. It is zero within a municipality and enters the models as $\log(1+d_{rj})$, together with indicators for the same municipality and the same province. Centroid distance is a proxy for spatial friction rather than route distance or travel time, and setting it to zero within a municipality understates internal commutes in large cities.

The capital region comprises Seoul, Incheon, and Gyeonggi. Indicators for a capital-region residence and a capital-region workplace define four residence-workplace market cells.

Work24 provides youth job openings by municipality and half-year, and the openings enter distance-decayed measures of job access (Section 4.5). A workplace municipality is a high-rent center in a half-year if its standardized rent lies in the top quartile of workplaces, and a high-opening center if its youth openings lie in the top quartile. Expected wages predicted from a Mincer-type earnings regression (Mincer, 1974) on the survey's wage sample sort workers into quartiles, and 2024 SGIS housing statistics give static measures of small-unit housing supply. These variables serve only the mechanism tests.

3.5 Descriptive facts

Figure 2 shows how far residence and workplace separate. About half of the estimation sample lives and works in the same municipality, and commutes beyond 30 km are uncommon (panel b). Among all employed persons aged 15–34 in 2024H2, the largest cross-province flows connect Gyeonggi, Incheon, and Seoul (panel c).

Rent levels differ sharply across regions (Table 2, Panel B). Measured in 10,000 won per month at 2024 prices, the median standardized rent across rent-eligible municipality-half-year cells is 63.8 in Seoul, 45.9 in Gyeonggi, and 42.4 in Incheon, against 35.5 outside the capital region.

The four market cells differ greatly in size. Workers who live and work in the capital region carry 57.44 percent of the survey weight, and workers who live and work outside it carry 41.77 percent. Commuting across the capital boundary is rare. Capital-region residents with noncapital workplaces carry 0.44 percent of the weight, and noncapital residents with capital-region workplaces carry 0.35 percent. A few large noncapital cities near the capital boundary account for most of the latter cell, which matters for the capital-workplace contrast in Section 5.3. Appendix Figure A1 maps the workplace distribution.

4 Empirical Strategy

4.1 Workplace-conditioned residential choice

Let g index worker types, t half-years, and j the observed workplace municipality, and let $r \in \mathcal{R}_t$ denote a residence municipality with rent support in half-year t. A worker of type g employed in j obtains utility $U_{grjt} = V_{grjt} + \varepsilon_{grjt}$ from residence r, with

$$\begin{aligned} V_{grjt} = \lambda_{gtj} + \beta_H h_{rt} + \beta_D \log(1 + d_{rj}) + \beta_S \mathbf{1}\{r = j\} \\ + \beta_P \mathbf{1}\{p(r) = p(j)\} + \boldsymbol{\gamma}' \mathbf{A}_{rt} + \rho_{p(r)}. \end{aligned} \tag{1}$$

Here h_{rt} is log standardized rent, $p(\cdot)$ maps a municipality to its province, $\mathbf{A}_{rt}$ collects the four amenity indices with coefficient vector $\boldsymbol{\gamma}$, and $\rho_{p(r)}$ is a residence-province effect with Seoul omitted. The coefficients β_D, β_S, and β_P describe commuting frictions through distance, the same-municipality indicator, and the same-province indicator. The choice-group effect λ_{gtj} absorbs everything that is common to the residences available to a type-g worker employed in j in half-year t.

The model is conditional on workplace. Expected wages, job openings, and every other attribute of workplace j are constant across the residences in a choice group, so they cancel from the conditional likelihood. The model therefore yields no workplace-choice coefficients, and none are reported. The design compares residences for workers who already hold a job in a given place. It is not a joint model of employment and workplace choice, and it does not treat the workplace as exogenous, since workers may choose jobs and homes together.

If ε_{grjt} is independently and identically distributed type-I extreme value, the probability of residence r takes the conditional logit form (McFadden, 1974)

$$P_{grjt} = \frac{\exp(V_{grjt})}{\sum_{q \in \mathcal{R}_t} \exp(V_{gqjt})}. \tag{2}$$

Let y_{grjt} be the rescaled survey-weighted number of workers in choice group (g, t, j) who live in r. The grouped conditional log likelihood for the parameter vector $\boldsymbol{\theta}$ is

$$\ell(\boldsymbol{\theta}) = \sum_g \sum_t \sum_j \sum_{r \in \mathcal{R}_t} y_{grjt} \log P_{grjt}(\boldsymbol{\theta}). \tag{3}$$

Workers in the same cell face identical attributes, so the grouped likelihood equals the individual likelihood up to a constant (Guimaraes et al., 2003). When residential coefficients are common to all types, the counts can be summed over g within (t, j, r), which is the exact collapse used for the more demanding specifications.

The coefficient β_H is a conditional rent gradient. Holding the workplace and measured attributes fixed, it describes how the relative odds of choosing a residence change with the residence's standardized rent. It is not a demand elasticity. Rent is an equilibrium price that also reflects unobserved quality, productivity, housing supply, and sorting, which motivates the decomposition in Section 4.3.

4.2 Full and sampled choice sets

The main specifications use every rent-eligible residence, so each choice group contains 207–212 alternatives and no sampling correction is required. As a diagnostic, sampled choice sets keep every residence with a positive observed flow and draw zero-flow residences within strata defined by distance band and residence size. If a zero-flow residence enters the sampled set $\mathcal{R}^S_{gtj}$ with probability π_{rgtj}, the sampled probability includes an offset (McFadden, 1978),

$$P^S_{grjt} = \frac{\exp(V_{grjt} + a_{rgtj})}{\sum_{q \in \mathcal{R}^S_{gtj}} \exp(V_{gqjt} + a_{qgtj})}, \qquad a_{rgtj} = -\log \pi_{rgtj}, \tag{4}$$

with $\pi_{rgtj} = 1$ for residences with positive flows. The diagnostics draw 3, 6, or 12 zero-flow residences per stratum under five random seeds and compare estimates with and without the offset against the full-choice estimate (Figure 4 and Appendix Table A2). Full choice sets remove any dependence on sampled alternatives, but they do not relax the independence of irrelevant alternatives assumption of the conditional logit. Appendix Table A9 reports alternative functional forms for distance and the estimates obtained with the deposit-inclusive rent index of Section 3.3.

4.3 Within-between decomposition and municipality fixed effects

A pooled rent gradient combines two kinds of variation: changes in a municipality's rent over time and persistent differences in rent across municipalities. For a stable residence geography r observed in the set $\mathcal{T}_r$ of T_r half-years, let

$$\bar{h}_r = \frac{1}{T_r} \sum_{t \in \mathcal{T}_r} h_{rt}, \qquad h^W_{rt} = h_{rt} - \bar{h}_r, \tag{5}$$

and define the same components for each amenity index k,

$$\bar{A}_{k,r} = \frac{1}{T_r} \sum_{t \in \mathcal{T}_r} A_{k,rt}, \qquad A^W_{k,rt} = A_{k,rt} - \bar{A}_{k,r}. \tag{6}$$

After the exact collapse, the Mundlak specification is

$$\begin{aligned} V^M_{rjt} = \lambda_{tj} + \beta_W h^W_{rt} + \beta_B \bar{h}_r + \beta_D \log(1 + d_{rj}) + \beta_S \mathbf{1}\{r = j\} \\ + \beta_P \mathbf{1}\{p(r) = p(j)\} + \sum_k \left(\gamma_{Wk} A^W_{k,rt} + \gamma_{Bk} \bar{A}_{k,r} \right) + \rho_{p(r)}, \end{aligned} \tag{7}$$

where λ_{tj} is a half-year-by-workplace effect. The within coefficient β_W compares a municipality with its own average rent over the period. The between coefficient β_B compares municipalities with different persistent rent levels and therefore captures between-municipality sorting. The difference $\beta_W - \beta_B$ is tested directly. Neither coefficient is causal. Time-varying local shocks can move rent and residential choice together, and persistent rent differences can reflect sorting and omitted local quality.

Two specifications control more strongly for time-invariant differences between municipalities. The first replaces the province effects and the between components with a residence-municipality effect α_r,

$$\begin{aligned} V^{FE}_{rjt} = \lambda_{tj} + \alpha_r + \beta_W h^W_{rt} + \boldsymbol{\gamma}'_W \mathbf{A}^W_{rt} \\ + \beta_D \log(1 + d_{rj}) + \beta_S \mathbf{1}\{r = j\} + \beta_P \mathbf{1}\{p(r) = p(j)\}, \end{aligned} \tag{8}$$

where $\mathbf{A}^W_{rt}$ collects the within components of the amenity indices. The second adds a municipality-specific linear trend $\tau_r t$ to the municipality effects and keeps rent and the commuting terms but not the within amenity terms. With seven half-years, this trend model is a demanding sensitivity check rather than a dynamic model. Further timing checks start the sample in 2022H1, replace current rent with the previous half-year's rent, or use the year-on-year change in log rent (Appendix Table A4). These checks are neither instruments nor tests for pretrends.

4.4 Capital-region residence and workplace markets

Let C_r and C_j indicate a residence and a workplace in the capital region. A first extension of equation (7) lets the within and between rent slopes differ between capital-region and noncapital residences for all workers, with between slopes $\beta_B^{C_r=1}$ and $\beta_B^{C_r=0}$. The market model instead lets both slopes differ across the four cells $m(r,j) = (C_r, C_j)$:

$$\begin{aligned} V_{rjt}^{2\times 2} = \lambda_{tj} + \sum_{m} \left[\beta_W^m h_{rt}^W + \beta_B^m \bar{h}_r\right] \mathbf{1}\{m(r,j) = m\} \\ + \mathbf{Z}_{rjt}' \boldsymbol{\delta} + \rho_{p(r)} + C_j \psi_{p(r)}, \end{aligned} \tag{9}$$

where $\mathbf{Z}_{rjt}$ contains the distance terms and the within and between amenity components with coefficient vector $\boldsymbol{\delta}$, and $\beta_B^{(C_r,C_j)}$ denotes the between slope of cell $m = (C_r, C_j)$. The interaction $C_j\psi_{p(r)}$ lets the attractiveness of each residence province differ between capital-region and noncapital workplace markets. The two contrasts of interest are

$$\Delta_B = \beta_B^{C_r=1} - \beta_B^{C_r=0}, \qquad \Delta_B^{C_j=1} = \beta_B^{(1,1)} - \beta_B^{(0,1)}. \tag{10}$$

The pooled contrast Δ_B compares the between slopes of capital-region and noncapital residences in the first extension. The capital-workplace contrast $\Delta_B^{C_j=1}$ makes the same comparison within equation (9) for workers employed in the capital region. Each contrast is a difference between two slopes. A large contrast can come from a steep negative slope for capital-region residences, from a positive slope for noncapital residences, or from both, so the component slopes are always reported. The contrast for noncapital workplaces and the difference-in-differences across workplace markets are reported as well.

Few noncapital residents work in the capital region, so the noncapital-residence slope in that market rests on a small number of cities near the capital boundary. A sensitivity check removes from every choice set the five, and then the ten, noncapital municipalities that carry the largest weight among noncapital residents employed in the capital region, and it re-estimates both contrasts. Province-specific slopes for Seoul, Incheon, and Gyeonggi, the omission of each province, alternative rent measures, housing-type-specific rents, and a central common-support sample complete the robustness checks for the market structure.

4.5 Commuting trade-offs and alternative mechanisms

For residence r in half-year t and a decay parameter κ, distance-decayed job access is

$$JA_{rt}(\kappa) = \log\left[1 + \sum_j O_{jt} \exp\left(-\frac{d_{rj}}{\kappa}\right)\right], \qquad \kappa \in \{30, 60\} \text{ km}, \tag{11}$$

where O_{jt} is youth openings at workplace j. The index is standardized within half-year, and a 30 km variant replaces openings with the survey-weighted stock of employed young workers. The access indices enter equation (7) as within and between components. The test asks whether high-rent municipalities are chosen because they give access to a larger surrounding job market.

The commuting test lets the rent and distance coefficients differ between workers at high-rent workplace centers and other workers. In observed flows, a residence is cheaper when its log rent is below that of the workplace, and it is far when the centroid distance exceeds 30 km. The shares of cheaper-near and cheaper-far residences are then compared across the two groups, and a parallel test uses high-opening centers. Center status, rent, and distance are determined jointly by metropolitan structure, so the test describes a spatial pattern rather than a causal mediation.

Three predeclared alternatives complete the mechanism tests. The rent slope may differ by quartile of expected wages or by a static measure of small-unit housing supply. In a third test, a residual rent purged of access, amenities, and housing supply replaces log rent in the between component. Appendix Table A5 reports every result, including the null estimates.

4.6 Inference

The likelihood treats rescaled survey-weighted flows as choice totals. To allow dependence across spatial alternatives, the covariance matrix is clustered two ways, by stable residence geography and by stable workplace geography (Cameron et al., 2011). With $\widehat{V}_R$, $\widehat{V}_J$, and $\widehat{V}_{RJ}$ denoting the sandwich covariance matrices clustered by residence, by workplace, and by residence-workplace pair, the two-way estimator is

$$\widehat{V}_{\text{2way}} = \widehat{V}_R + \widehat{V}_J - \widehat{V}_{RJ}, \tag{12}$$

with finite-cluster corrections. This estimator is not guaranteed to be positive semidefinite. In the sparse municipality fixed-effect specifications, some variances of nuisance parameters are negative, but none of the reported coefficients is affected.

The number of geographic clusters is limited, so the main estimands also receive wild-score intervals based on 999 Rademacher draws (Kline and Santos, 2012), computed separately over residence blocks and over workplace blocks. The routine is not available for the sparse fixed-effect specifications, for which only two-way clustered inference is reported. Model checks compare observed and fitted residence shares and re-estimate the model without one half-year at a time.

4.7 Employment-selection sensitivity

The choice sample contains only wage workers, because a nonemployed person has no observed workplace to condition on. To assess sensitivity to selection on observed characteristics, the analysis starts from all persons aged 19–34 and distinguishes nonemployment, other employment, and wage employment. After residence attributes are linked, the employment-state model contains 409,456 observations. A survey-weighted multinomial logit with nonemployment as the base state gives

$$\Pr(S_{it} = s \mid \mathbf{X}_{it}) = \frac{\exp(\mathbf{X}_{it}'\boldsymbol{\eta}_s)}{1 + \sum_{k\in\{\text{other, wage}\}} \exp(\mathbf{X}_{it}'\boldsymbol{\eta}_k)}, \qquad s \in \{\text{other}, \text{wage}\}, \tag{13}$$

where S_{it} is the employment state of person i in half-year t, $\boldsymbol{\eta}_s$ is the coefficient vector for state s, and $\mathbf{X}_{it}$ contains standardized residence rent, 30 km opening access, the four amenity indices, sex, age band, education, school status, marital status, and half-year and residence-province effects. Let $\widehat{p}_i$ be the fitted probability of wage employment for wage worker i and $\bar{p}$ the survey-weighted share of wage employment. The stabilized weight is

$$w_i^{IPW} = w_i^A \, \frac{\bar{p}}{\widehat{p}_i}, \tag{14}$$

where w_i^A is the rescaled analysis weight. The baseline clips the stabilized factor at its weighted 1st and 99th percentiles and rescales the weights to the sample size. Clipping at the 5th and 95th percentiles and overlap weights proportional to $w_i^A(1-\widehat{p}_i)$ are alternatives. The reweighted records are collapsed again, and equations (7) and (9) are re-estimated.

The reweighting moves the observed characteristics of wage workers toward those of all persons aged 19–34. It cannot give a nonemployed person a potential workplace or address selection on unobservables. It is therefore a sensitivity analysis of sample scope, not an extension of the outcome to all young people (Appendix Table A6).

4.8 Interpretation boundaries

The design narrows a broad spatial matching problem to a transparent conditional question: among employed wage workers aged 19–34 attached to a given workplace in a given half-year, how do measured residence attributes covary with where they live? Five limits follow from this framing.

First, standardized rent is an equilibrium price. Municipality effects and trends absorb time-invariant and linear differences, but unobserved quality and endogenous supply may remain. The estimates describe conditional gradients and spatial sorting, and this is not a causal analysis of rent. Second, the index is the market price of a standardized small unit. It does not measure a worker's own contract, rent-to-income ratio, or housing quality. Third, seven half-years are too few for dynamic or pretrend claims, and the analysis uses no exogenous shock to housing costs. Fourth, centroid distance measures commuting imperfectly, especially within large municipalities. Fifth, the conditional logit imposes independence of irrelevant alternatives across residences. Flexible capital-region slopes and province effects capture observed heterogeneity, while unobserved taste heterogeneity is not modeled.

All conclusions therefore concern employed wage workers aged 19–34 whose residence has rent support. Differences by age, sex, and education are a predeclared auxiliary test reported in Appendix Table A8. Appendix Table A7 records the reproducibility gates and the claim registry that governs the wording of the results.

5 Main Results

This section reports the results in the order of the argument. Section 5.1 establishes the conditional rent gradient in the full choice set and shows that it does not depend on sampled alternatives, seeds, survey rounds, or the inference method. Section 5.2 decomposes the gradient into within- and between-municipality variation and locates it in persistent differences between municipalities. Section 5.3 shows that the between-municipality gradient is confined to capital-region residences and separates that result from a boundary result for workers employed in the capital region. Section 5.4 documents the commuting pattern around expensive workplace centers as a descriptive mechanism. Every estimate is a conditional association among employed wage workers aged 19–34 with a given workplace; the interpretation boundaries of Section 4.8 apply throughout.

5.1 The full-choice rent gradient

Table 3 reports the full-choice benchmark of equations (1)–(3), estimated over every rent-eligible residence municipality with residence-province fixed effects. The log-rent coefficient in Panel A, column 1, is -1.153 (0.263). Because the regressor is log rent, one unit corresponds to a 2.7-fold difference in standardized rent. On the scale that separates neighboring municipalities, a residence whose standardized rent is 10 percent higher has an estimated -10.4 percent change in its relative odds of being chosen, holding the workplace and the other measured attributes fixed.

The commuting terms confirm that spatial friction shapes the residential choice of a worker with a given workplace, which is what the workplace-conditioned design requires. The coefficient on log distance is -1.919, and the same-municipality and same-province indicators are precisely estimated. These terms and the amenity indices are controls; they are not interpreted individually, because the same-municipality indicator is defined jointly with a zero centroid distance (Section 3.4).

The benchmark is not an artifact of the choice set or of any single survey round. Figure 4 and Appendix Table A2 report three checks. Across five random seeds, sampled choice sets with 12 zero-flow alternatives per stratum and the McFadden offset give a mean coefficient of -1.155, which is 0.17 percent away from the full-choice value; omitting the offset moves the estimate to -1.049, so the correction matters even though the full choice set makes it unnecessary (panel a). Re-estimating the full-choice model with one half-year omitted yields coefficients between -1.201 and -1.120 (panel b). Predicted municipality residence shares track observed shares with a Spearman correlation of 0.978 in sample and between 0.975 and 0.979 for the omitted half-years (panel c).

Panel C of Table 3 compares inference methods. The two-way clustered 95 percent interval for the pooled coefficient excludes zero, and so do the wild-score intervals over 999 residence-block and workplace-block draws, whose lower bounds are -1.513 and -1.569. The same holds for the capital-minus-noncapital contrast introduced next, whose wild-score intervals are $[-2.654, -1.293]$ and $[-2.703, -1.187]$.

Column (2) of Panel A interacts log rent with a capital-region residence indicator in the same pooled specification, without separating within- and between-municipality variation. The implied slope is -1.446 for capital-region residences and 0.534 ($p = 0.105$) for noncapital residences, a difference of -1.979 (Panel B). Section 5.3 takes up this contrast with persistent rent levels and shows what it does and does not depend on.

The gradient is a conditional association, not a causal demand elasticity. Standardized rent is an equilibrium price that reflects unobserved local quality, productivity, and housing supply as well as housing cost (Section 4.8). The decomposition that follows asks which kind of rent variation carries the gradient, not what a change in rent would cause.

5.2 Where the gradient comes from: within versus between

Does the pooled gradient come from changes in a municipality's own rent over time or from persistent differences in rent levels across municipalities? Table 4 Panel A reports the Mundlak decomposition of equation (7). The within-municipality coefficient is 0.080 ($p = 0.273$) and the between-municipality coefficient is -1.254 ($p < 0.001$); their difference of 1.334 is precisely estimated ($p < 0.001$). Figure 5 plots the two coefficients with their two-way clustered and residence-block wild-score intervals (panel a).

Table 3: Full-choice benchmark and capital-region rent gradient

Panel A. Coefficients

	(1) Pooled	(2) Capital interaction
Log rent	−1.153 (0.263)	0.534 (0.329)
Log rent x capital residence		−1.979 (0.428)
Log (1 + distance)	−1.919 (0.057)	−1.926 (0.057)
Same municipality	−2.431 (0.190)	−2.445 (0.190)
Same province	0.642 (0.087)	0.633 (0.087)
Medical index	−0.056 (0.065)	−0.070 (0.062)
Cultural index	−0.614 (0.072)	−0.575 (0.069)
Education index	0.440 (0.067)	0.431 (0.064)
Transport index	0.057 (0.055)	0.028 (0.045)
Persons	224,862	224,862
Choice groups	21,379	21,379
Choice-set rows	4,498,851	4,498,851
Alternatives per group	207–212	207–212
Residence-province fixed effects	Yes	Yes

Panel B. Implied slopes from column (2)

	Estimate (SE)	p-value
Noncapital residences	0.534 (0.329)	0.105
Capital residences	−1.446 (0.275)	<0.001
Capital minus noncapital	−1.979 (0.428)	<0.001

Panel C. Inference comparison (95% intervals)

	Two-way clustered	Residence-block wild	Workplace-block wild
Pooled log rent	[−1.668, −0.638]	[−1.513, −0.756]	[−1.569, −0.729]
Capital minus noncapital	[−2.819, −1.140]	[−2.654, −1.293]	[−2.703, −1.187]

Notes: Workplace-conditioned conditional logit over every rent-eligible residence municipality, with residence-province fixed effects. Coefficients are conditional associations, not causal effects. Standard errors in parentheses are clustered two-way by stable residence and workplace geography. Wild-score intervals use 999 Rademacher draws over residence or workplace blocks. The capital region is Seoul, Incheon, and Gyeonggi.

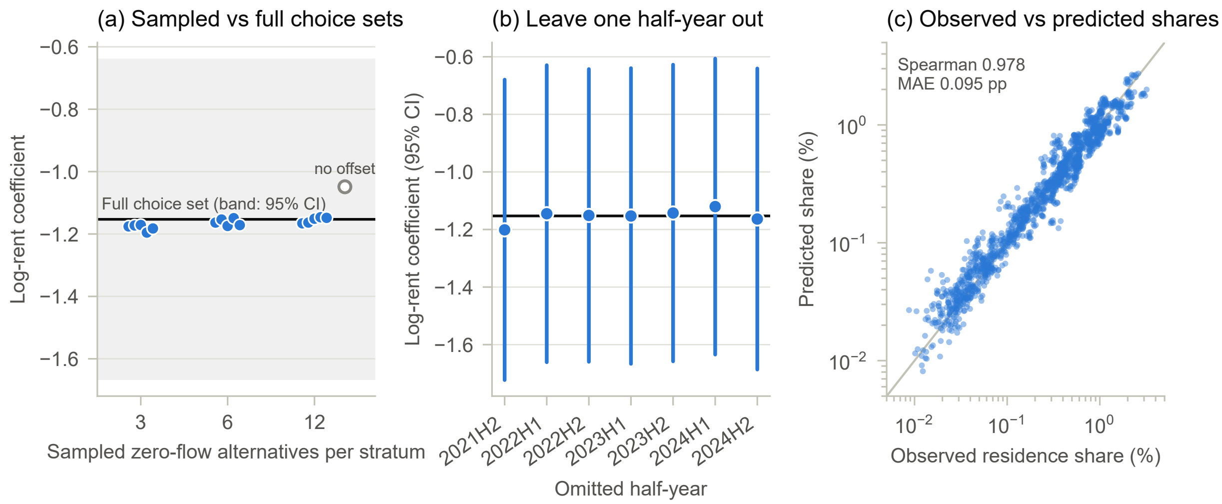

Figure 4: Full-choice stability and predictive validation

Notes: Panel (a) shows five random seeds for each number of sampled zero-flow alternatives, with the McFadden offset, against the full-choice estimate and its two-way clustered 95% interval; the hollow marker omits the offset. Panel (b) re-estimates the full-choice model dropping one half-year at a time. Panel (c) plots municipality residence shares by half-year on log scales.

A municipality whose seven-half-year average rent is 10 percent higher has an estimated −11.3 percent change in the relative odds of being chosen, whereas a 10 percent deviation of a municipality's rent from its own average is associated with no comparable change. The wild-score intervals confirm the split. For the between coefficient they are [−1.646, −0.855] over residence blocks and [−1.690, −0.813] over workplace blocks; for the within coefficient they are [−0.037, 0.212] and [−0.067, 0.226]; both intervals for the difference exclude zero.

Two specifications that control more strictly for time-invariant differences between municipalities reach the same conclusion. Replacing the province effects and the between components with a municipality fixed effect gives a within coefficient of 0.103 ($p = 0.203$), and adding municipality-specific linear trends gives 0.041 ($p = 0.508$) (Panel B). Because these models use sparse municipality effects, only two-way clustered inference is available for them.

Panel C previews the capital-workplace contrast of Section 5.3 under alternative timing. The contrast is −5.512 with region-by-half-year effects, −5.469 when the sample starts in 2022H1, and −4.576 with rent lagged one half-year, but only 0.344 ($p = 0.642$) when current rent is replaced by its year-on-year change (Appendix Table A4). Lagged and differenced rents are timing checks, not instruments. They show that the pattern rides on persistent rent levels rather than on recent rent movements.

The pooled gradient of Section 5.1 therefore arises mainly from persistent between-municipality rent differences: it is spatial sorting across municipalities, not a response to short-run rent changes. A within-municipality coefficient statistically indistinguishable from zero does not establish that rent changes within a municipality are unrelated to residential choice. Seven half-years provide little within-municipality variation, and time-varying local shocks can move rent and residential choice together, so the within coefficient is best read as uninformative in this design rather than as evidence that rent changes do not matter (Section 4.3).

Table 4: Within-between decomposition, municipality fixed effects, and timing

Panel A. Pooled Mundlak model

	Estimate (SE)	p-value	Residence-block wild 95%	Workplace-block wild 95%
Within-municipality rent	0.080 (0.073)	0.273	[−0.037, 0.212]	[−0.067, 0.226]
Between-municipality rent	−1.254 (0.281)	<0.001	[−1.646, −0.855]	[−1.690, −0.813]
Within minus between	1.334 (0.295)	<0.001	[0.901, 1.757]	[0.868, 1.794]

Panel B. Municipality fixed effects

	Estimate (SE)	p-value	Residence-block wild 95%	Workplace-block wild 95%
Within rent, municipality FE	0.103 (0.081)	0.203	n.a.	n.a.
Within rent, municipality FE and trends	0.041 (0.062)	0.508	n.a.	n.a.

Panel C. Timing, capital-region workplaces

	Capital residences	Noncapital residences	Difference
Baseline (current rent, between)	−1.515 (0.282)	4.016 (1.187)	−5.531 (1.206)
Adding region-by-half-year effects	−1.515 (0.282)	3.997 (1.184)	−5.512 (1.203)
Sample from 2022H1	−1.542 (0.281)	3.926 (1.215)	−5.469 (1.237)
One-half-year lagged rent	−1.452 (0.267)	3.123 (0.799)	−4.576 (0.823)
Year-on-year rent change	0.326 (0.339)	−0.019 (0.751)	0.344 (0.741)

Notes: Panel A decomposes log rent into the municipality mean over seven half-years (between) and the deviation from that mean (within), with the same decomposition for the four amenity indices. Wild-score intervals for Panel A were computed in S1 on a rebuilt panel that reproduces every frozen clustered standard error. Panel B models use sparse municipality fixed effects, for which the wild-score routine is not available. Panel C reports between-municipality slopes for workers employed in the capital region; lagged and year-on-year specifications replace the rent measure and are not instruments. The difference column inherits the dependence on commuter cities adjacent to the capital region documented in Table 5. Standard errors in parentheses are clustered two-way by stable residence and workplace geography. Estimates are conditional associations, not causal effects.

5.3 Capital-region market structure

Where is the between-municipality gradient located? Table 5 Panel A lets the between slope differ between capital-region and noncapital residences for all workers, as in the first extension

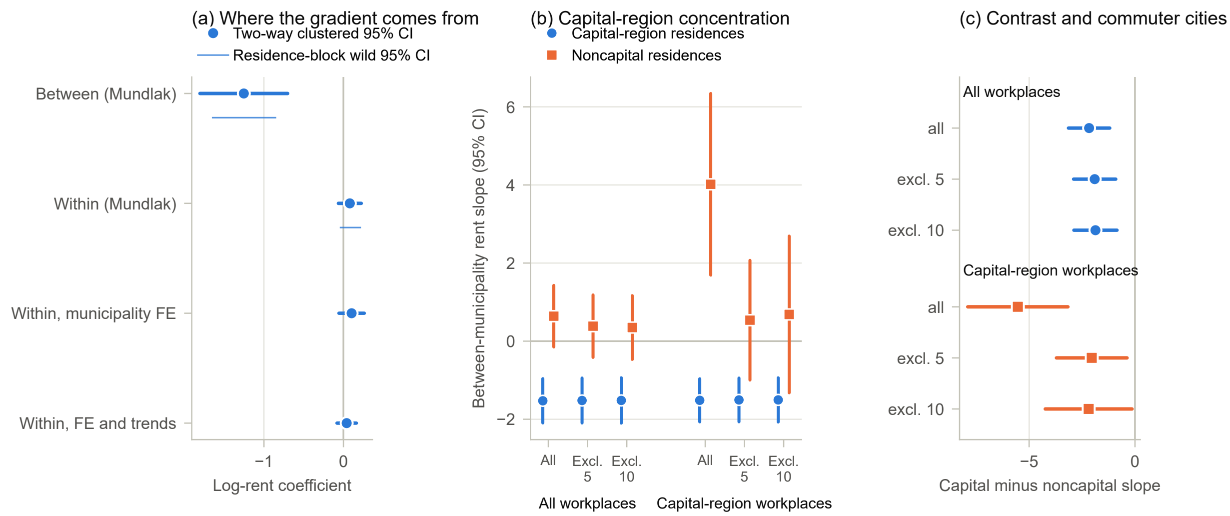

Figure 5: Spatial origin of the rent gradient

Notes: Panel (a) reports log-rent coefficients from the pooled Mundlak model and from municipality fixed-effect models; the thin line is the residence-block wild-score interval, which is not available for fixed-effect models. Panels (b) and (c) report between-municipality slopes; 'Excl. 5' and 'Excl. 10' remove the noncapital residences carrying the largest weight among noncapital residents employed in the capital region (Cheonan, Asan, Wonju, Cheongju, Chuncheon, then Sejong, two Daejeon districts, Eumseong, Jincheon) from every choice set. Intervals are two-way clustered 95% intervals; estimates are conditional associations.

of equation (9). The slope is -1.527 for capital-region residences and 0.639 ($p = 0.112$) for noncapital residences, a contrast Δ_B of -2.166 ($p < 0.001$). The between-municipality gradient is therefore a capital-region phenomenon. Across noncapital residences, persistent rent differences are not detectably related to residential choice.

This contrast does not depend on the few noncapital cities that dominate cross-boundary commuting, which are identified below. Removing the five largest of them from every choice set leaves -1.520 for capital-region residences, 0.384 ($p = 0.348$) for noncapital residences, and a contrast of -1.903 ($p < 0.001$); removing ten leaves a contrast of -1.867 ($p < 0.001$) with a noncapital slope of 0.349 ($p = 0.402$). The pooled interaction of Section 5.1 behaves the same way, with a contrast of -1.760 after the five-city exclusion and -1.736 after the ten-city exclusion. Figure 5 shows the slopes (panel b) and the contrasts (panel c) before and after the exclusions.

Panel B narrows the comparison to workers employed in the capital region, using the residence-by-workplace market model of equation (9). Among these workers, the between slope is -1.515 for capital-region residences and 4.016 for noncapital residences, so the contrast $\Delta_B^{C_j=1}$ of equation (10) is -5.531 (1.206; $p < 0.001$), with a residence-block wild-score interval of $[-8.054, -3.073]$. The difference-in-differences against noncapital workplaces is -4.564 ($p < 0.001$), with a residence-block interval of $[-7.251, -1.838]$. Read on its own, this contrast would suggest a far sharper ordering by rent among capital-region workers than Panel A shows for all workers.

That reading does not survive inspection of the cell that produces it. Noncapital residents employed in the capital region carry only 0.35 percent of the survey weight, and five municipalities account for 71.3 percent of that cell: Cheonan (33.4 percent), Asan (18.5), Wonju (9.2), Cheongju (5.2), and Chuncheon (5.0); the ten largest account for 85.3 percent (Appendix Figure A3). Removing the five from every choice set leaves the capital-region slope at −1.505 but reduces the noncapital slope to 0.537 ($p = 0.492$), whose wild-score intervals, [−0.891, 1.948] and [−1.053, 2.135], now include zero. The contrast falls to −2.042 ($p = 0.016$), with wild-score intervals of [−3.499, −0.555] and [−3.731, −0.345] that still exclude zero; removing ten cities gives −2.188 ($p = 0.036$), and removing Cheonan alone already reduces it to −3.605. The difference-in-differences falls to −1.264 ($p = 0.185$) after the five-city exclusion and −1.445 ($p = 0.208$) after the ten-city exclusion, and is then no longer distinguishable from zero.

Two further contrasts are reported with the same weight. For workers employed outside the capital region, the capital-minus-noncapital contrast is −0.968 ($p = 0.107$); its residence-block wild-score interval, [−2.237, 0.335], includes zero, while its workplace-block interval, [−1.871, −0.079], excludes it, so its significance depends on which geographic dimension is treated as the clustering unit. The within-municipality contrast for capital-region workplaces is −0.447 ($p = 0.479$) in the baseline, −1.604 ($p = 0.040$) after the five-city exclusion, and −1.125 ($p = 0.175$) after the ten-city exclusion; it is imprecise and unstable across exclusions, and no conclusion rests on it. The noncapital-residence slope among capital-region workers is also nearly collinear with the workplace-market-by-residence-province intercepts of equation (9) and is not interpreted as a freestanding estimate.

The five municipalities behind the capital-workplace contrast are not commuter towns oriented toward Seoul; they are large noncapital labor markets close to the capital boundary (Figure 6; Appendix Table A3, Panels E and F). Among noncapital rent-eligible municipalities, their youth job-stock percentiles run from the 85th to the 100th (Cheongju 100th, Cheonan 99th, Asan 95th, Wonju 88th, Chuncheon 85th). Between 73 and 92 percent of their resident wage workers are employed within their own city, and only 0.9 to 8.0 percent commute to the capital region. The nearest capital-region municipality lies 23 to 50 kilometers away, and their standardized rent ranks relatively high among noncapital municipalities, at the 59th to the 87th percentile. The largest destinations of their capital-region commuters are municipalities just across the boundary in outer Gyeonggi: from Cheonan and Asan to Pyeongtaek, Anseong, and Hwaseong; from Wonju to Yeoju and Icheon; from Cheongju to Icheon; and from Chuncheon to Gapyeong, followed by the Gangnam and Songpa districts of

Seoul. Among these commuters, 40 percent travel 40 kilometers or less, compared with 12 percent among the remaining noncapital residents employed in the capital region, whose mean commuting distance is 121 kilometers. The survey counts are small, with 427 commuters for the five cities combined, including 26 for Cheongju and 43 for Chuncheon, so industry composition is described only in aggregate; manufacturing accounts for 21 percent of the five-city commuters.

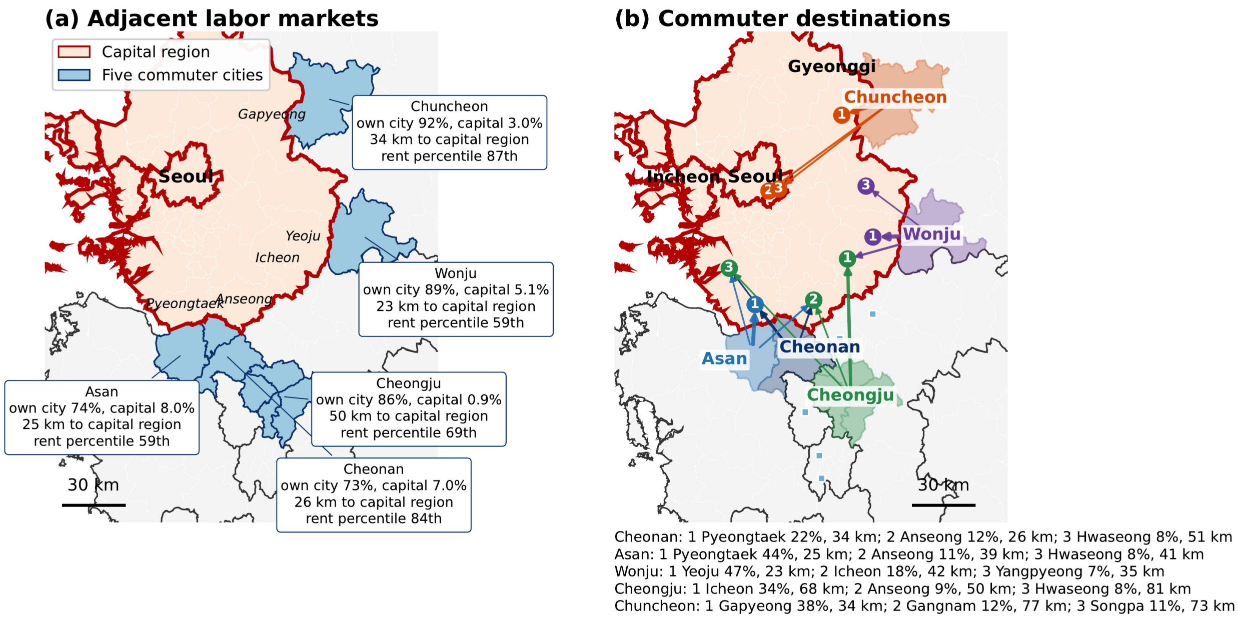

Figure 6: The five commuter cities at the capital-region boundary

Notes: Both panels zoom on the capital-region boundary; Seoul is drawn as one unit. Panel (a) shades the five noncapital municipalities that carry the largest survey weight among noncapital residents employed in the capital region and reports, for each, the share of its resident wage workers employed in the city itself (own city) and in the capital region (capital), the centroid distance to the nearest capital-region municipality, and the city's rent percentile among noncapital rent-eligible municipalities (Appendix Table A3, Panel E). Panel (b) draws an arrow from each city to the three largest destinations of its capital-region commuters; the list below the panel gives each destination's share of the city's capital-region commuters and the mean centroid distance of those commutes, and the arrow width is proportional to the share. Squares mark the commuter cities ranked sixth to tenth. Italic labels in panel (a) name the Gyeonggi municipalities that receive most of these commuters.

The positive noncapital-residence slope in Panel B therefore does not show capital-workplace commuters avoiding high rent within the capital region. It reflects commuting across the administrative capital boundary between adjacent labor markets, in which a large and relatively expensive noncapital city is chosen as a residence over closer and cheaper alternatives. It should not be read as evidence that these commuters prefer expensive housing.

Panels A and B thus say two different things, and the paper keeps them apart. The between-municipality gradient is confined to capital-region residences, and that concentration survives the removal of the commuter cities. Economically, the capital region combines the concentration of jobs with a rent structure that ranks its municipalities steeply, and the residential choices of workers attached to it are ordered by those persistent rent levels; outside the capital region, where rent levels differ less across municipalities (Section 3.5), no such ordering is detected. Among capital-region workers, the additional contrast created by the positive noncapital slope is a boundary result that depends on a handful of adjacent cities. It is reported only together with the exclusions and is not read as showing that capital-region residents are simply more sensitive to rent than noncapital residents. The estimate for workers employed outside the capital region is imprecise. Section 6.2 shows that the capital-region slope is not driven by Seoul alone and keeps its sign on the central common rent support, which retains only about one-fifth of capital-region residence cells.

5.4 Commuting trade-offs around expensive workplace centers

If the between-municipality gradient reflects sorting around expensive workplaces, workers attached to the most expensive workplaces should show it most clearly, together with the commuting that such sorting requires. Table 5 Panel C classifies each workplace municipality, within its half-year, as a high-rent center if its standardized rent is in the top quartile of workplaces and, separately, as a high-opening center if its youth job openings are in the top quartile.[2]

Workers employed at high-rent centers have a rent slope of −1.739, compared with −0.238 ($p = 0.441$) for workers at other workplaces, a difference of −1.501 (0.357; $p < 0.001$); their distance slope differs by only −0.047 ($p = 0.395$) (Figure 7, panel a). Panel D and Figure 7, panel b, describe the observed flows behind this contrast. Workers at high-rent centers commute 10.09 kilometers on average, compared with 7.07 kilometers for other workers. Of the workers at high-rent centers, 48.5 percent live in a municipality whose standardized rent is below that of their workplace, compared with 11.2 percent of other workers; 43.4 percent live in such a cheaper municipality within 30 kilometers, compared with 9.9 percent; and 5.12 percent live in a cheaper municipality more than 30 kilometers away, compared with 1.32 percent. Only 37.5 percent live in the same municipality as their workplace, compared with 66.2 percent.

High-opening centers show a different pattern, reported with the same weight. The rent-slope difference for high-opening centers is 0.075 ($p = 0.788$), and the rent slopes of center and other workplaces are both negative and similar in size; the distance slope at high-opening centers is 0.132 ($p = 0.002$) less negative than elsewhere. Density of job openings

[2]High-rent centers are concentrated in the capital region. Of the 401 center-half-year cells, 83.0 percent are capital-region municipalities; 93.4 percent of the workers employed at high-rent centers have capital-region workplaces, and 56.3 percent have workplaces in Seoul.

Table 5: Capital-region concentration, commuter cities, and commuting trade-offs

Panel A. Capital versus noncapital residences, all workplaces

	All residences	Excluding five	Excluding ten
Capital residences	−1.527 (0.289)	−1.520 (0.294)	−1.518 (0.296)
Noncapital residences	0.639 (0.402)	0.384 (0.408)	0.349 (0.417)
Capital minus noncapital	−2.166 (0.496)	−1.903 (0.507)	−1.867 (0.519)

Panel B. By workplace market

	All residences	Excluding five	Excluding ten
Capital workplaces: capital residences	−1.515 (0.282)	−1.505 (0.286)	−1.504 (0.289)
Capital workplaces: noncapital residences	4.016 (1.187)	0.537 (0.782)	0.684 (1.024)
Capital workplaces: difference	−5.531 (1.206)	−2.042 (0.850)	−2.188 (1.045)
Noncapital workplaces: difference	−0.968 (0.600)	−0.778 (0.604)	−0.744 (0.621)
Difference-in-differences	−4.564 (1.234)	−1.264 (0.953)	−1.445 (1.147)
Capital workplaces: within-municipality difference	−0.447 (0.633)	−1.604 (0.780)	−1.125 (0.830)

Panel C. Workplace centers

	High-rent centers	High-opening centers
Rent slope, other workplaces	−0.238 (0.309)	−1.236 (0.242)
Rent slope, center workplaces	−1.739 (0.265)	−1.161 (0.306)
Rent slope difference	−1.501 (0.357)	0.075 (0.279)
Distance slope difference	−0.047 (0.055)	0.132 (0.043)

Panel D. Observed commuting, high-rent centers

	Center workplaces	Other workplaces
Mean distance (km)	10.09	7.07
Cheaper residence (%)	48.5	11.2
Cheaper and within 30 km (%)	43.4	9.9
Cheaper and beyond 30 km (%)	5.12	1.32
Same municipality (%)	37.5	66.2

Notes: Panels A and B report between-municipality rent slopes from Mundlak models. The excluded commuter cities are the residences carrying the largest weight among noncapital residents employed in the capital region: Cheonan, Asan, Wonju, Cheongju, and Chuncheon (five), plus Sejong, Daejeon Seo-gu, Daejeon Yuseong-gu, Eumseong, and Jincheon (ten); they are removed from every choice set. The last row of Panel B reports the within-municipality contrast for capital-region workplaces. Panel C allows rent and distance slopes to differ for workplaces in the top quartile of standardized rent or youth openings within each half-year. Panel D describes observed flows; a residence is cheaper when its log rent is below the workplace's. Standard errors in parentheses are clustered two-way by stable residence and workplace geography. Estimates are conditional associations, not causal effects.

by itself does not produce the pattern; a high rent level at the workplace does.

Workers tied to an expensive workplace center therefore live cheaper but farther: they show a steeper rent gradient, more often live in a municipality whose rent is below that of their workplace, and commute longer than workers at other workplaces. Because high-rent centers lie mostly in the capital region, this is the worker-level counterpart of the capital-region concentration in Section 5.3: the steep ordering by persistent rent among capital-region residences is accompanied by residence in cheaper neighboring municipalities and longer commutes. The pattern is descriptive, not a causal mediation analysis. Center status, residential rent, and commuting distance are jointly determined by metropolitan structure (Section 4.5), and Section 6.3 shows that job access, expected wages, and static housing supply do not account for the gradient differences.

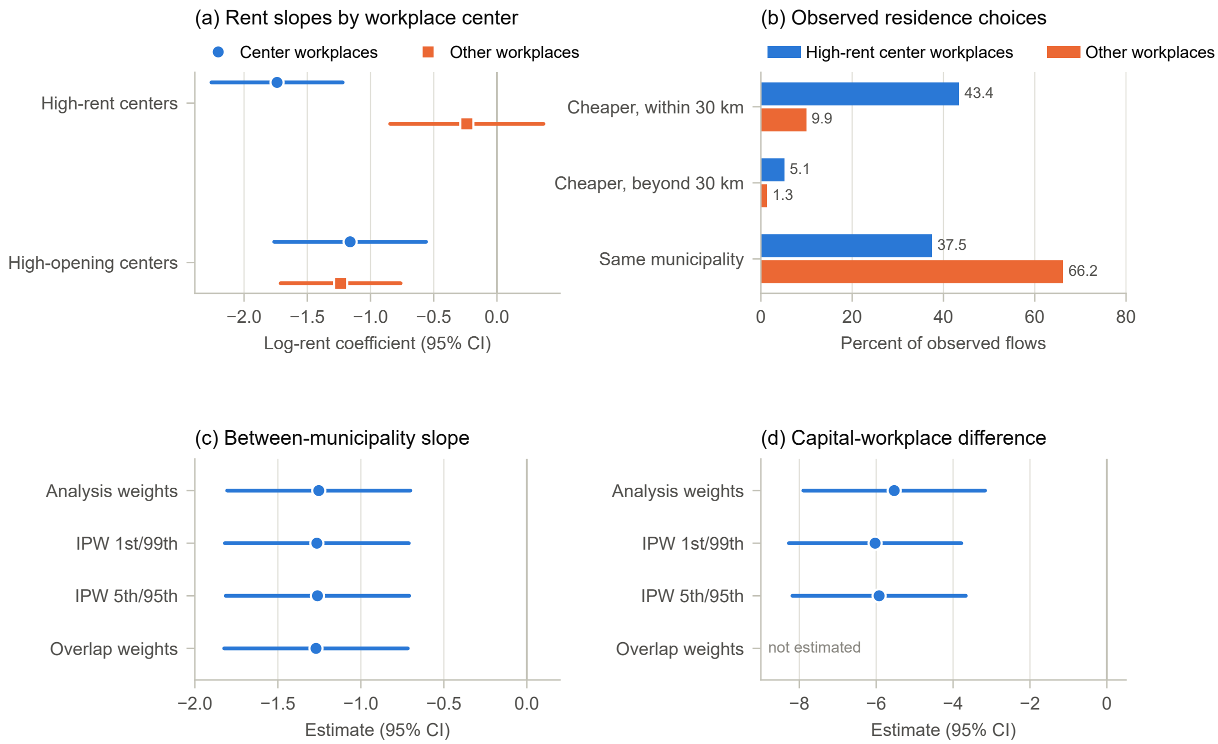

Figure 7: Commuting trade-offs and employment-selection sensitivity

Notes: Panel (a) allows rent slopes to differ for workplaces in the top quartile of standardized rent or youth openings within each half-year. Panel (b) describes observed flows for the rent-center definition; a residence is cheaper when its log rent is below the workplace's. Panels (c) and (d) reweight wage workers toward all persons aged 19–34 on observed characteristics; the overlap weight was not estimated for the market-structure model. Patterns are descriptive and not causal mediation.

6 Robustness and Employment-Selection Sensitivity

Table 6 collects the checks that qualify the results of Section 5, and Appendix Tables A2, A3, A5, A6, and A8 give the details. The checks sort the results by the confidence each warrants.

6.1 Rent definitions and housing types

Table 6 Panel A re-estimates the residence-by-workplace market model with pooled rent slopes under alternative rent measures. For workers employed in the capital region, the capital-minus-noncapital slope difference is −4.276 with the baseline index, −3.180 with monthly-rent contracts only, −4.541 under a fixed 4 percent deposit conversion, and −4.285 for a 60 m^2 standard unit. The last is not independent evidence: the hedonic floor-area coefficient is common across municipalities, so the two size standards have a spatial correlation of 1.000. These differences inherit the commuter-city dependence of Section 5.3; Panel A shows that the rent measure does not produce them. For workers employed outside the capital region, the difference stays imprecise under every measure, between −0.585 and −0.932, against −0.719 ($p = 0.110$) in the baseline.

A deposit-inclusive variant of the index also admits *jeonse* leases, converted at the same market rate; deposit-only leases are 40.0 percent of the 5,166,950 contracts behind it in the rent-eligible cells, and its log rent correlates 0.920 with the baseline across those cells (Appendix Table A1, Panel F). With this index the capital-workplace difference is −2.290 (0.552) and the noncapital-workplace difference −0.461 ($p = 0.058$); the pooled slope is −1.045 (0.236) and the capital-minus-noncapital difference −1.588 (0.337) (Appendix Table A9, Panel B). The gradients are smaller with deposits included but keep their sign.

Housing-type-specific rent indices give the same ordering (Appendix Table A3, Panel C). The capital-region slope is negative for every type, from −0.843 for apartments to −2.194 for single- and multi-family houses; no noncapital slope is distinguishable from zero, and the difference is least precise for row houses (−0.887; $p = 0.092$).

6.2 Geography and common support

Panel B of Table 6 lets the pooled rent slope differ by province. It is −1.977 for Seoul residences, −1.240 for Gyeonggi, and −0.626 ($p = 0.259$) for Incheon, against 0.486 ($p = 0.147$) for noncapital residences, so the gradient is not Seoul alone. Removing each province from every choice set in turn leaves the capital-minus-noncapital difference, −1.979 in the pooled specification of Section 5.1, at −1.539 without Seoul, −2.098 without Incheon, and −2.210 without Gyeonggi (Appendix Table A3, Panel B).

Table 6: Robustness and employment-selection sensitivity

Panel A. Rent measures

	Capital workplaces	Noncapital workplaces
Market-rate conversion (baseline)	−4.276 (0.944)	−0.719 (0.450)
Monthly rent only	−3.180 (0.689)	−0.585 (0.558)
Fixed 4 percent conversion	−4.541 (1.047)	−0.932 (0.517)
60 m^2 standard unit	−4.285 (0.945)	−0.726 (0.449)
Deposit-inclusive conversion	−2.290 (0.552)	−0.461 (0.243)

Panel B. Geography and support

	Estimate (SE)	p-value
Seoul residence slope	−1.977 (0.407)	<0.001
Incheon residence slope	−0.626 (0.554)	0.259
Gyeonggi residence slope	−1.240 (0.364)	<0.001
Noncapital residence slope	0.486 (0.335)	0.147
Capital minus noncapital, excluding Seoul	−1.539 (0.539)	0.004
Capital minus noncapital, excluding Incheon	−2.098 (0.438)	<0.001
Capital minus noncapital, excluding Gyeonggi	−2.210 (0.504)	<0.001
Capital minus noncapital, central common support	−2.564 (1.162)	0.027

Panel C. Employment-selection reweighting

	Between-municipality slope	Capital-workplace difference
Analysis weights	−1.254 (0.281)	−5.531 (1.206)
IPW trimmed at 1st/99th	−1.265 (0.283)	−6.028 (1.144)
IPW trimmed at 5th/95th	−1.261 (0.282)	−5.924 (1.151)
Overlap weights	−1.270 (0.282)	n.a.

Table 6: Robustness and employment-selection sensitivity (continued)

Panel D. Weight diagnostics		
	Effective sample size	Max absolute SMD
Analysis weights	172,221	0.333
IPW trimmed at 1st/99th	135,009	0.035
IPW trimmed at 5th/95th	149,235	0.080
Overlap weights	155,081	0.120

Notes: Panel A re-estimates the market-structure model with alternative rent measures; the 60 m^2 standard is not independent evidence because the hedonic area coefficient is common nationally, so its spatial correlation with the 40 m^2 measure is 1.000; the deposit-inclusive row uses an index that also admits deposit-only (jeonse) leases converted at the same market rate, with the choice sets unchanged (Appendix Tables A1 and S-A9). Panel B central common support retains 19.7–24.2 percent of capital residence cells per half-year. Panel C reweights wage workers toward all persons aged 19–34 on observed characteristics; it does not recover potential workplaces of nonworkers or correct selection on unobservables. The capital-workplace differences in Panels A and C come from the residence-by-workplace market model and inherit the dependence on commuter cities adjacent to the capital region documented in Table 5. Standard errors in parentheses are clustered two-way by stable residence and workplace geography. Estimates are conditional associations, not causal effects.

Common support is the more serious limit, since capital-region municipalities occupy the upper rent range. Keeping in each half-year only the alternatives whose log rent lies within the 5th–95th percentile range shared by the two groups gives a difference of −2.564 ($p = 0.027$), with component slopes of −1.972 ($p = 0.069$) and 0.593 ($p = 0.329$), on a sample that retains only 19.7–24.2 percent of capital residence cells per half-year (Panel D). The contrast keeps its sign on common support, but a result resting on about one-fifth of capital-region cells does not show that it holds across the full rent distribution, so its external validity is limited.

Lowering the minimum contract count per cell from 20 to 10 or 5 enlarges the choice sets to 220–225 and 225–228 municipalities and leaves the pooled slope at −1.150 and the capital-minus-noncapital difference at −1.976 and −1.968, against −1.153 and −1.979 at the baseline threshold (Appendix Table A2, Panel D).

6.3 Mechanisms that are not supported

Three explanations were tested and are not supported (Appendix Table A5). The first is that expensive municipalities are chosen because they offer access to more jobs. Adding a standardized job-access measure to equation (7) as within and between terms leaves the between rent slope between −1.240 and −1.251, while the between component of access is −0.011 ($p = 0.947$) for youth openings within 30 km, −0.029 ($p = 0.873$) within 60 km, and −0.035 ($p = 0.851$) for the employed-youth stock within 30 km (Panel A). With access held constant, the capital-workplace contrast in the market model is −5.625 (Panel B). Persistent differences in municipal job access do not account for the gradient or its capital-region concentration.

The second is that workers with lower earning capacity are more sensitive to rent. Letting the pooled slope differ by expected-wage quartile gives slopes between −1.071 and −1.304; the lowest quartile, at −1.304, differs from the highest, at −1.114, by −0.190 ($p = 0.435$) (Panel C). The third is that static small-unit housing supply shapes the gradient. The between slope is −1.183 where the supply constraint is low and −1.260 where it is high, a difference of −0.077 ($p = 0.829$), and the residual of log rent after projection on access, amenities, and supply has a between slope of −1.226 (Panel D); the residual slope is not an estimate of amenity capitalization. Municipal aggregates and a single 2024 supply cross-section give these tests limited power against subtler versions of the mechanisms.

6.4 Employment-selection sensitivity

Wage workers are 54.3 to 58.1 percent of persons aged 19–34 across the seven half-years (Appendix Table A6, Panel A), so the choice sample omits about two-fifths of the age group. The employment-state model of equation (13), fitted to 409,456 persons, has an AUC of 0.725 for wage employment. In the wage-versus-nonemployment equation, standardized residence rent has a coefficient of 0.004 ($p = 0.855$) and 30 km opening access 0.030 ($p = 0.434$) (Panel B); residence attributes barely separate the two groups. The predicted wage-employment probability of wage workers runs from 0.139 to 0.786 between its 1st and 99th percentiles. Clipping the weights there reduces the effective sample size from 172,221 to 135,009, or 78.4 percent, and the largest standardized mean difference from 0.333 to 0.035; the 5th–95th percentile and overlap weights retain 149,235 and 155,081 effective observations with maximum differences of 0.080 and 0.120 (Table 6, Panel D).

The spatial estimates barely move (Panel C; Figure 7, panels c and d). The between-municipality slope is −1.265, −1.261, and −1.270 under the three schemes, against −1.254 with analysis weights, and under the 1st–99th percentile weights its residence-block wild-score interval, [−1.655, −0.870], excludes zero. The capital-workplace contrast is −6.028 and −5.924 against −5.531, with an interval of [−8.316, −3.782]. Reweighting on observed employment selection therefore does not overturn the core spatial gradients.

The exercise is narrower than a correction: the weights balance observed characteristics only, cannot give a nonemployed person a potential workplace, and rest on no exclusion restriction, so selection on unobservables is untouched. The results remain statements about employed wage workers aged 19–34, which is why the paper's scope is defined that way.

6.5 Demographic heterogeneity

Differences by age, sex, and education were predeclared as uncertain in sign and are reported regardless of significance (Appendix Table A8). With one interaction at a time, the pooled slope is negative and precisely estimated in every group, between −1.069 and −1.291; the differences are −0.107 ($p = 0.226$) for workers aged 19–29 relative to 30–34, −0.168 ($p = 0.086$) for women relative to men, and 0.197 ($p = 0.201$) for four-year college graduates. Jointly, the three interactions give a Wald statistic of 3.192 on 3 degrees of freedom ($p = 0.363$), or 3.056 ($p = 0.383$) with the capital-region residence interaction added, which leaves the capital-minus-noncapital difference at −1.988 ($p < 0.001$). Differences in the rent gradient by age, sex, and education are therefore not statistically distinguishable from zero.

6.6 What the robustness checks do not establish

The checks leave the five boundaries of Section 4.8 in place and bear on two of them. On independence of irrelevant alternatives, removing the alternatives of Seoul, Incheon, or Gyeonggi in turn leaves the capital-region contrast close to its full-choice value (Section 6.2), the stability across subsets that tests of the assumption examine, but unobserved taste heterogeneity is not modeled. On distance, bins or a quadratic in log distance leave the pooled slope at −1.021 and −1.125 and the capital-minus-noncapital difference at −1.767 and −1.862 (Appendix Table A9, Panel A), but centroid distance still measures commuting imperfectly within large municipalities. The other boundaries are untouched: rent remains an equilibrium price rather than an exogenous treatment, so none of the estimates is a causal elasticity; seven half-years support no dynamic or pretrend claims; and the index is the market rent of a standardized small unit, not the worker's own contract or rent-to-income burden. One limit is added by the design itself: the workplace is conditioned on, not chosen, so the joint choice of workplace and residence is not identified. Every estimate remains a conditional association among employed wage workers aged 19–34 with a given workplace.

7 Conclusion

This paper asked where young wage workers live, given where they work. Linking seven half-years of survey microdata on employed wage workers aged 19–34 to a standardized rent index built from administrative lease records, it modeled the residence municipality as a conditional logit over every rent-eligible municipality, with the observed workplace held fixed. Three results answer the questions posed at the outset. The conditional log-rent gradient is negative (−1.153) and does not depend on the choice set, the survey round, or the inference method. It arises from persistent differences in rent levels between municipalities (-1.254) rather than from changes in a municipality's own rent (0.080). And it is confined to capital-region residences: the capital-minus-noncapital contrast is -2.166, remains -1.903 after the noncapital cities that dominate cross-boundary commuting are removed from every choice set, and is accompanied, among workers at high-rent workplace centers, by a steeper gradient and by residence in cheaper and more distant municipalities. Reweighting on observed employment selection does not overturn these gradients, and job access, expected wages, static housing supply, and demographic composition do not account for them.

The economic reading is one of persistent spatial sorting rather than a short-run price response. Young workers attached to the capital region's jobs are ordered across municipalities by durable differences in rent levels, and the ordering is not visible in half-year movements of rent within a municipality or across noncapital residences, where rent levels differ less. What produces the pattern is the joint concentration, in one metropolitan region, of the jobs that draw young workers and the highest rents in the country: workers hold the job, and the residential margin is where the cost of the region is absorbed, partly through commuting from cheaper municipalities. The finding speaks to other economies organized around a single dominant metropolis, where a national rent coefficient would average a steep metropolitan gradient with a flat one elsewhere.

The implication for policy is a design principle rather than a program. Because the gradient is a feature of the capital region's workplace markets and is expressed through residence in cheaper municipalities and longer commutes, measures aimed at young workers' housing are unlikely to work in isolation from small-unit supply near expensive workplace centers, commuting access, and the spatial distribution of jobs. The paper does not evaluate any such measure. It estimates conditional associations, not responses to an exogenous change in rent, and it says nothing about welfare or about the housing costs that individual workers actually pay.

Five limits bound these conclusions, and they point to the next steps. The design uses no exogenous variation in housing costs, so rent remains an equilibrium price that also carries unobserved quality, supply, and sorting. The choice sample is conditional on employment; the reweighting balances observed characteristics only and cannot give the nonemployed a workplace, so the results describe employed wage workers rather than all young people. Seven half-years permit no dynamic or pretrend claims. The conditional logit imposes independence of irrelevant alternatives across residences, and centroid distance measures commuting coarsely within large municipalities. The rent index is a market rent for a standardized small unit; a deposit-inclusive variant leaves the gradients in place with smaller magnitudes. Future work can relax the substitution structure with nested or mixed logit specifications, follow workers across survey rounds where panel identifiers allow, model the choice between deposit and monthly-rent tenure, and use discrete supply events, such as large-scale completions near workplace centers, to move from conditional associations toward identified responses.

Data availability

The Regional Employment Survey microdata were obtained from the Microdata Integrated Service (MDIS) of Statistics Korea and are available to researchers under its access conditions; the author cannot redistribute them. Lease contracts from the Real Transaction Management System (RTMS), youth job openings from Work24, local service indicators from the e-Local Indicators of KOSIS, and boundaries and centroids from SGIS are public. The municipality-half-year rent index, the derived choice-set tables, the number ledger that documents every reported figure, and the estimation code will be deposited in a public repository at publication.

Declarations

Funding. This study received no specific funding.
Competing interests. The author declares no competing interests.

References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The economics of density: Evidence from the Berlin Wall. *Econometrica*, 83(6):2127–2189.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.

Choi, H.-C. and Lee, S.-B. (2017). Analysis of local youth labor market with special emphasis on out/in flow of college graduate [in Korean]. *Journal of Industrial Economics and Business*, 30(2):505–545.

Choi, H.-j. (2022). The effects of social and cultural infrastructure on the regional migration of college graduates [in Korean]. *Labor Policy Research*, 22(2):97–125.

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.

Ganong, P. and Shoag, D. (2017). Why has regional income convergence in the U.S. declined? *Journal of Urban Economics*, 102:76–90.

Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature*, 47(4):983–1028.

Guimaraes, P., Figueiredo, O., and Woodward, D. (2003). A tractable approach to the firm location decision problem. *Review of Economics and Statistics*, 85(1):201–204.

Han, J. and Yoon, C. (2021). A study on interregional migration of youth: Focused on college enrollment [in Korean]. Technical report, Korea Development Institute, Policy Study 2021-09.

Heblich, S., Redding, S. J., and Sturm, D. M. (2020). The making of the modern metropolis: Evidence from London. *Quarterly Journal of Economics*, 135(4):2059–2133.

Hsieh, C.-T. and Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39.

Kennan, J. and Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251.

Kline, P. and Santos, A. (2012). A score based approach to wild bootstrap inference. *Journal of Econometric Methods*, 1(1):23–41.

Manning, A. and Petrongolo, B. (2017). How local are labor markets? Evidence from a spatial job search model. *American Economic Review*, 107(10):2877–2907.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record*, 673:72–77.

Mincer, J. A. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research, New York.

Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–3890.

Moretti, E. (2011). Local labor markets. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics, Volume 4B*, pages 1237–1313. Elsevier, Amsterdam.

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6):1257–1278.

Saks, R. E. (2008). Job creation and housing construction: Constraints on metropolitan area employment growth. *Journal of Urban Economics*, 64(1):178–195.

Shim, J. and Kim, E. (2012). Analysis of occupational mobility of college graduates [in Korean]. *The Korea Spatial Planning Review*, 75:37–51.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(5):80–93.

Appendix: Additional Tables

Table A1: Administrative crosswalk and rent-measure construction

Panel A. MDIS to RTMS municipality bridge		
	Bridge rows	MDIS municipalities
SGIS boundary code to legal-district code	279	258
Crosswalk with current RTMS query code	44	44
Current RTMS query code override	7	2
Explicit special-city bridge	5	5
Incheon district reorganization bridges	5	3
All rules	340	312

Table A1: Administrative crosswalk and rent-measure construction (continued)

Panel B. Pooled hedonic model of log monthly-equivalent rent			
	Market-rate conversion (baseline)	Monthly rent only	Fixed 4 percent conversion
Floor area (centered)	0.190 (0.000)	0.152 (0.000)	0.205 (0.000)
Floor area squared	0.037 (0.000)	0.045 (0.000)	0.034 (0.000)
Building age (centered)	−0.092 (0.000)	−0.013 (0.000)	−0.119 (0.000)
Building age squared	0.009 (0.000)	−0.003 (0.000)	0.013 (0.000)
Building age missing	−0.040 (0.001)	−0.033 (0.001)	−0.033 (0.001)
Floor (centered)	0.076 (0.001)	0.080 (0.001)	0.076 (0.001)
Floor squared	−0.010 (0.000)	−0.015 (0.001)	−0.010 (0.000)
Floor missing	0.084 (0.000)	0.156 (0.000)	0.058 (0.000)
Officetel	0.409 (0.001)	0.575 (0.001)	0.340 (0.001)
Row house	0.213 (0.001)	0.270 (0.001)	0.179 (0.001)
Single- or multi-family house	0.084 (0.000)	0.156 (0.000)	0.058 (0.000)
Within R^2	0.270	0.107	0.365
Contracts, 2017H1–2024H2	5,159,595	5,159,595	5,159,595
Municipality-half-year fixed effects	4,104	4,104	4,104

Panel C. Housing-type-specific hedonic models					
	Contracts	Municipality-half-year cells	Within R^2	Floor-area coefficient	$\log(60\ m^2)$ minus $\log(40\ m^2)$
Apartment	1,879,624	4,020	0.361	0.226	0.714
Officetel	731,478	2,967	0.390	0.155	0.196
Row house	657,031	3,820	0.167	0.101	0.160
Single- or multi-family house	1,891,462	4,090	0.217	0.115	0.204
Pooled, fixed type mix					0.527

Table A1: Administrative crosswalk and rent-measure construction (continued)

Panel D. Validation against raw medians and R-ONE		
	Municipality-half-year cells	Correlation
Standardized vs raw median, market-rate conversion	4,104	0.849
Standardized vs raw median, fixed 4 percent conversion	4,104	0.831
Standardized vs raw median, monthly rent only	4,104	0.865
MOLIT standardized (fixed 4 percent) vs R-ONE deposit-adjusted rent	1,862	0.852

Panel E. Contract coverage in the analysis window				
	Municipalities with contracts	Rent-eligible municipalities	Contracts, all cells	Contracts, eligible cells
2021H2	229	211	394,874	394,658
2022H1	229	211	462,188	462,001
2022H2	229	212	426,063	425,844
2023H1	229	207	455,261	455,015
2023H2	229	209	441,610	441,348
2024H1	229	212	478,124	477,935
2024H2	229	211	445,755	445,548
All half-years	n.a.	n.a.	3,103,875	3,102,349

Table A1: Administrative crosswalk and rent-measure construction (continued)

Panel F. Deposit-inclusive rent variant, rent-eligible cells 2021H2–2024H2	
	Value
Contracts behind the monthly-rent index	3,102,349
Contracts behind the deposit-inclusive index	5,166,950
Deposit-only (jeonse) share of contracts (%)	40.0
Pearson correlation of log standardized rent	0.920
Spearman correlation of log standardized rent	0.897
Mean within-half-year Pearson correlation	0.942
Mean log gap, deposit-inclusive minus monthly-rent index	−0.183

Notes: Panel A maps the survey's municipality codes (SGIS boundary codes) to the legal-district codes used by the RTMS lease registry; a municipality can enter several rules when districts were merged, split, or recoded during the period. Panel B regresses the log monthly-equivalent rent of each contract on floor area, building age, floor, housing type, and month-within-half-year dummies with municipality-by-half-year fixed effects; the standard errors are homoskedastic OLS errors reported for completeness. The standardized rent is the fixed effect evaluated at 40 m^2 and a fixed housing-type mix. Panel C estimates the same model separately by housing type; the 60 m^2 standard shifts every cell's log rent by the constant in the last column, so it carries no independent spatial variation. Panel D correlates the standardized measure with raw municipality medians and with the R-ONE deposit-adjusted rent where both exist. The hedonic sample covers 2017H1–2024H2; Panel E reports the analysis window 2021H2–2024H2, in which a municipality-half-year cell is rent-eligible with at least 20 contracts. Panel F compares the baseline index, built from monthly-rent contracts, with a deposit-inclusive variant that also admits deposit-only (jeonse) leases converted at the same market rate; the correlations are over the rent-eligible municipality-half-year cells of the analysis window.

Table A2: Alternative sampling, predictive diagnostics, and contract-support thresholds

Panel A. Full and sampled choice sets

	Log-rent coefficient (SE)	95% CI	p-value
Full choice set	−1.153 (0.263)	[−1.668, −0.638]	<0.001
3 sampled alternatives per stratum, seed 1	−1.176 (0.261)	[−1.688, −0.663]	<0.001
3 sampled alternatives per stratum, seed 2	−1.173 (0.264)	[−1.689, −0.656]	<0.001
3 sampled alternatives per stratum, seed 3	−1.171 (0.263)	[−1.686, −0.656]	<0.001
3 sampled alternatives per stratum, seed 4	−1.195 (0.266)	[−1.716, −0.675]	<0.001
3 sampled alternatives per stratum, seed 5	−1.182 (0.267)	[−1.706, −0.659]	<0.001
6 sampled alternatives per stratum, seed 1	−1.163 (0.263)	[−1.680, −0.647]	<0.001
6 sampled alternatives per stratum, seed 2	−1.154 (0.261)	[−1.664, −0.643]	<0.001
6 sampled alternatives per stratum, seed 3	−1.174 (0.262)	[−1.687, −0.661]	<0.001
6 sampled alternatives per stratum, seed 4	−1.149 (0.262)	[−1.663, −0.635]	<0.001
6 sampled alternatives per stratum, seed 5	−1.171 (0.260)	[−1.681, −0.662]	<0.001
12 sampled alternatives per stratum, seed 1	−1.166 (0.263)	[−1.681, −0.651]	<0.001
12 sampled alternatives per stratum, seed 2	−1.163 (0.261)	[−1.674, −0.652]	<0.001
12 sampled alternatives per stratum, seed 3	−1.151 (0.261)	[−1.662, −0.640]	<0.001
12 sampled alternatives per stratum, seed 4	−1.146 (0.262)	[−1.659, −0.633]	<0.001
12 sampled alternatives per stratum, seed 5	−1.148 (0.263)	[−1.664, −0.632]	<0.001
12 sampled alternatives per stratum, seed 1, no offset correction	−1.049 (0.255)	[−1.548, −0.549]	<0.001

Panel B. Leave one half-year out

	Log-rent coefficient (SE)	95% CI	Holdout MAE (pp)	Holdout Spearman
Omit 2021H2	−1.201 (0.266)	[−1.723, −0.679]	0.098	0.975
Omit 2022H1	−1.145 (0.263)	[−1.661, −0.629]	0.095	0.979
Omit 2022H2	−1.151 (0.259)	[−1.660, −0.643]	0.097	0.978
Omit 2023H1	−1.153 (0.262)	[−1.667, −0.640]	0.098	0.979
Omit 2023H2	−1.143 (0.263)	[−1.658, −0.628]	0.097	0.977
Omit 2024H1	−1.120 (0.262)	[−1.635, −0.606]	0.091	0.979
Omit 2024H2	−1.164 (0.267)	[−1.686, −0.641]	0.093	0.977

Table A2: Alternative sampling, predictive diagnostics, and contract-support thresholds (continued)

Panel C. Observed versus predicted residence shares, full choice set

	Region-half-year cells	MAE (pp)	RMSE (pp)	Pearson	Spearman
2021H2	211	0.098	0.177	0.939	0.975
2022H1	211	0.094	0.177	0.941	0.979
2022H2	212	0.097	0.186	0.935	0.978
2023H1	207	0.098	0.179	0.938	0.979
2023H2	209	0.097	0.173	0.941	0.977
2024H1	212	0.090	0.166	0.946	0.978
2024H2	211	0.092	0.167	0.945	0.977
All half-years	1,473	0.095	0.175	0.941	0.978

Panel D. Minimum contracts per residence cell

	Alternatives per half-year	Persons	Excluded, no rent cell	Rent slope (SE)	Capital slope (SE)	Noncapital slope (SE)	Difference (SE)
20 contracts (baseline)	207–212	224,862	5,063	−1.153 (0.263)	−1.446 (0.275)	0.534 (0.329)	−1.979 (0.428)
10 contracts	220–225	228,253	1,672	−1.150 (0.263)	−1.455 (0.276)	0.521 (0.318)	−1.976 (0.421)
5 contracts	225–228	229,314	611	−1.150 (0.263)	−1.457 (0.276)	0.511 (0.315)	−1.968 (0.420)

Notes: Panel A compares the full-choice log-rent coefficient with estimates from sampled choice sets that keep the chosen residence and every positive-flow alternative and add 3, 6, or 12 zero-flow alternatives per stratum with the McFadden sampling offset; the last row omits the offset. Panel B re-estimates the full-choice model dropping one half-year and predicts the residence shares of the omitted half-year. Panel C compares observed and predicted residence shares within region-half-year cells (percentage points). Panel D rebuilds the choice sets with lower minimum contract counts per residence cell and re-estimates the pooled slope and the capital-region interaction. Standard errors in parentheses are clustered two-way by stable residence and workplace geography unless stated otherwise. Estimates are conditional associations, not causal effects.

Table A3: Capital-region geography, housing types, common support, and commuter cities

Panel A. Rent slopes by capital-region province			
	Estimate (SE)	95% CI	p-value
Noncapital residences	0.486 (0.335)	[−0.171, 1.144]	0.147
Seoul	−1.977 (0.407)	[−2.774, −1.180]	<0.001
Incheon	−0.626 (0.554)	[−1.712, 0.460]	0.259
Gyeonggi	−1.240 (0.364)	[−1.953, −0.527]	<0.001
Seoul minus noncapital	−2.463 (0.514)	[−3.470, −1.456]	<0.001
Incheon minus noncapital	−1.112 (0.643)	[−2.372, 0.148]	0.084
Gyeonggi minus noncapital	−1.726 (0.496)	[−2.698, −0.755]	<0.001

Panel B. Residence-alternative subsets		
	Estimate (SE)	p-value
Three markets: capital residences	−1.450 (0.274)	<0.001
Three markets: noncapital metropolitan cities	0.022 (0.663)	0.973
Three markets: other noncapital residences	0.858 (0.335)	0.010
Excluding Seoul: capital residences	−1.344 (0.345)	<0.001
Excluding Seoul: noncapital residences	0.195 (0.418)	0.641
Excluding Seoul: difference	−1.539 (0.539)	0.004
Excluding Incheon: capital residences	−1.594 (0.280)	<0.001
Excluding Incheon: noncapital residences	0.504 (0.332)	0.130
Excluding Incheon: difference	−2.098 (0.438)	<0.001
Excluding Gyeonggi: capital residences	−1.429 (0.351)	<0.001
Excluding Gyeonggi: noncapital residences	0.781 (0.328)	0.017
Excluding Gyeonggi: difference	−2.210 (0.504)	<0.001
Noncapital residences only: pooled slope	0.493 (0.356)	0.166
Special and metropolitan cities only: capital residences	−1.078 (0.315)	<0.001
Special and metropolitan cities only: noncapital residences	0.429 (0.601)	0.475
Special and metropolitan cities only: difference	−1.507 (0.689)	0.029
Central common support: capital residences	−1.972 (1.083)	0.069
Central common support: noncapital residences	0.593 (0.607)	0.329
Central common support: difference	−2.564 (1.162)	0.027

Table A3: Capital-region geography, housing types, common support, and commuter cities (continued)

Panel C. Housing-type-specific standardized rent

	Noncapital slope (SE)	Capital slope (SE)	Difference (SE)	p-value of difference
Apartment	0.243 (0.252)	−0.843 (0.183)	−1.087 (0.321)	<0.001
Officetel	0.211 (0.567)	−1.027 (0.292)	−1.238 (0.632)	0.050
Row house	−0.612 (0.457)	−1.499 (0.280)	−0.887 (0.526)	0.092
Single- or multi-family house	−0.452 (0.497)	−2.194 (0.348)	−1.742 (0.585)	0.003

Panel D. Central common support by half-year

	Common lower bound	Common upper bound	Capital residence cells	Capital cells retained (%)	Noncapital cells retained (%)
2021H2	3.415	3.721	66	19.7	55.2
2022H1	3.510	3.782	66	21.2	56.6
2022H2	3.529	3.828	66	21.2	54.8
2023H1	3.565	3.829	66	21.2	53.2
2023H2	3.589	3.807	66	21.2	43.4
2024H1	3.620	3.804	66	24.2	39.7
2024H2	3.513	3.761	66	21.2	58.6

Panel E. Commuter cities adjacent to the capital region

	Share of cell (%)	Work in own city (%)	Work in capital region (%)	Distance to capital region (km)	Rent percentile	Job-stock percentile
Cheonan	33.4	73.0	7.0	26	84	99
Asan	18.5	73.7	8.0	25	59	95
Wonju	9.2	89.3	5.1	23	59	88
Cheongju	5.2	85.9	0.9	50	69	100
Chuncheon	5.0	91.7	3.0	34	87	85

Table A3: Capital-region geography, housing types, common support, and commuter cities (continued)

Panel F. Profiles of capital-region commuters and comparison groups

	Observations	Mean commute (km)	Within 40 km (%)	Seoul workplace (%)	Manufacturing (%)	Regular workers (%)
Five cities: commuters to capital-region workplaces	427	56.0	40.0	29.8	21.0	78.9
Other noncapital residents: commuters to capital-region workplaces	435	120.7	12.4	41.2	12.7	66.2
Five cities: residents working in their own city	7,937	0.0	100.0	0.0	28.6	78.9
Capital-region residents with capital-region workplaces	100,011	9.8	97.6	48.9	14.7	79.8

Notes: Panels A-D use the residence-by-workplace market model. Panel A lets the pooled rent slope differ for Seoul, Incheon, and Gyeonggi residences. Panel B re-estimates the capital-region contrast on subsets of residence alternatives; the central common-support sample keeps, in each half-year, the alternatives whose log rent lies within the 5th-95th percentile range shared by capital and noncapital residences. Panel C replaces the rent measure with housing-type-specific standardized rents. Panel D reports the common-support bounds (log rent) and the share of residence cells retained. Panel E describes the five noncapital cities that carry the largest weight among noncapital residents employed in the capital region; percentiles are among rent-eligible noncapital municipalities. Panel F profiles their capital-region commuters from the survey microdata; observation counts are unweighted and industry shares are descriptive. Standard errors in parentheses are clustered two-way by stable residence and workplace geography unless stated otherwise. Estimates are conditional associations, not causal effects.

Table A4: Timing and region-by-half-year specifications of the market model

Panel A. Between-municipality rent slopes

	Baseline	Region-by-half-year effects	Start 2022H1	Lagged rent	Year-on-year change
Capital residences, capital workplaces	−1.515 (0.282)	−1.515 (0.282)	−1.542 (0.281)	−1.452 (0.267)	0.326 (0.339)
Noncapital residences, capital workplaces	4.016 (1.187)	3.997 (1.184)	3.926 (1.215)	3.123 (0.799)	−0.019 (0.751)
Difference, capital workplaces	−5.531 (1.206)	−5.512 (1.203)	−5.469 (1.237)	−4.576 (0.823)	0.344 (0.741)
Capital residences, noncapital workplaces	−0.396 (0.466)	−0.397 (0.466)	−0.580 (0.493)	−0.376 (0.438)	0.641 (0.269)
Noncapital residences, noncapital workplaces	0.572 (0.400)	0.572 (0.400)	0.613 (0.415)	0.477 (0.344)	0.103 (0.214)
Difference, noncapital workplaces	−0.968 (0.600)	−0.968 (0.600)	−1.193 (0.636)	−0.853 (0.531)	0.539 (0.351)
Difference-in-differences	−4.564 (1.234)	−4.544 (1.231)	−4.276 (1.261)	−3.723 (0.913)	−0.194 (0.756)

Panel B. Within-municipality rent slopes

	Baseline	Region-by-half-year effects	Start 2022H1
Capital residences, capital workplaces	0.104 (0.117)	0.112 (0.115)	0.103 (0.116)
Noncapital residences, capital workplaces	0.551 (0.639)	0.117 (0.761)	−0.373 (0.505)
Difference, capital workplaces	−0.447 (0.633)	−0.005 (0.760)	0.476 (0.541)
Capital residences, noncapital workplaces	0.636 (0.594)	1.089 (0.861)	0.941 (0.569)
Noncapital residences, noncapital workplaces	0.049 (0.071)	0.036 (0.072)	−0.057 (0.077)
Difference, noncapital workplaces	0.587 (0.598)	1.053 (0.867)	0.999 (0.565)
Difference-in-differences	−1.034 (0.839)	−1.058 (0.785)	−0.523 (0.781)

Panel C. p-values of the two market contrasts

	Baseline	Region-by-half-year effects	Start 2022H1	Lagged rent	Year-on-year change
Difference, capital workplaces	<0.001	<0.001	<0.001	<0.001	0.642
Difference-in-differences	<0.001	<0.001	<0.001	<0.001	0.797

Table A4: Timing and region-by-half-year specifications of the market model (continued)

Panel D. Sample coverage by specification

	Sufficient-statistic rows	Choice groups	Residence alternatives	First half-year	Chosen weight retained (%)
Pooled Mundlak	335,839	1,596	290	2021H2	100.0
Capital residence interaction	335,839	1,596	290	2021H2	100.0
Residence-by-workplace market	335,839	1,596	290	2021H2	100.0
Region-by-half-year effects	335,839	1,596	290	2021H2	100.0
Start 2022H1	287,731	1,368	289	2022H1	86.1
Lagged rent	279,104	1,366	272	2022H1	86.0
Year-on-year change	229,111	1,141	265	2022H2	71.7
Municipality fixed effects	335,839	1,596	290	2021H2	100.0
Municipality effects by workplace market	335,839	1,596	290	2021H2	100.0
Municipality effects and trends	335,839	1,596	290	2021H2	100.0

Notes: All columns estimate the residence-by-workplace market model. Region-by-half-year effects add capital, Chungcheong, and other-region effects interacted with half-year. Start 2022H1 drops the first half-year. Lagged rent replaces current log rent with the previous half-year's value; year-on-year change replaces it with the change from the same half-year of the previous year. Lagged and differenced rents are timing checks, not instruments and not pretrend tests. The between slopes for noncapital residences with capital-region workplaces inherit the dependence on commuter cities documented in Table 5. Panel D reports the sufficient-statistic rows, choice groups, and the share of chosen weight retained relative to the full window. Standard errors in parentheses are clustered two-way by stable residence and workplace geography unless stated otherwise. Estimates are conditional associations, not causal effects.

Table A5: Job access, expected wage, and static supply mechanism tests

Panel A. Job access in the pooled Mundlak model		
	Estimate (SE)	p-value
Openings within 30 km: rent, within	0.083 (0.073)	0.253
Openings within 30 km: rent, between	−1.251 (0.286)	<0.001
Openings within 30 km: access, within	−0.158 (0.148)	0.284
Openings within 30 km: access, between	−0.011 (0.163)	0.947
Openings within 60 km: rent, within	0.082 (0.073)	0.263
Openings within 60 km: rent, between	−1.246 (0.286)	<0.001
Openings within 60 km: access, within	−0.088 (0.254)	0.730
Openings within 60 km: access, between	−0.029 (0.182)	0.873
Job stock within 30 km: rent, within	0.080 (0.073)	0.276
Job stock within 30 km: rent, between	−1.240 (0.288)	<0.001
Job stock within 30 km: access, within	−2.477 (0.684)	<0.001
Job stock within 30 km: access, between	−0.035 (0.186)	0.851

Panel B. Residence-by-workplace market model with 30 km openings access		
	Estimate (SE)	p-value
Between: capital residences, capital workplaces	−1.552 (0.294)	<0.001
Between: noncapital residences, capital workplaces	4.074 (1.130)	<0.001
Between: difference, capital workplaces	−5.625 (1.157)	<0.001
Between: difference, noncapital workplaces	−1.050 (0.605)	0.083
Between: difference-in-differences	−4.576 (1.183)	<0.001
Openings access, between	0.107 (0.167)	0.520
Openings access, within	−0.164 (0.148)	0.267

Table A5: Job access, expected wage, and static supply mechanism tests (continued)

Panel C. Rent slopes by expected-wage quartile		
	Estimate (SE)	p-value
Quartile 1 (lowest expected wage)	−1.304 (0.244)	<0.001
Quartile 2	−1.128 (0.238)	<0.001
Quartile 3	−1.071 (0.266)	<0.001
Quartile 4 (highest expected wage)	−1.114 (0.361)	0.002
Quartile 1 minus quartile 4	−0.190 (0.243)	0.435

Panel D. Static housing supply and residual rent		
	Estimate (SE)	p-value
Supply model: rent, within	0.086 (0.074)	0.248
Supply model: rent, between, at the mean constraint	−1.222 (0.284)	<0.001
Supply model: between slope, low constraint	−1.183 (0.294)	<0.001
Supply model: between slope, high constraint	−1.260 (0.372)	<0.001
Supply model: high minus low constraint	−0.077 (0.356)	0.829
Residual rent: within	0.095 (0.072)	0.191
Residual rent: between	−1.226 (0.287)	<0.001

Notes: Panel A adds a standardized job-access measure (youth openings or job stock within 30 or 60 km of the residence centroid, distance-weighted) to the pooled Mundlak model as within and between terms. Panel B adds 30 km openings access to the residence-by-workplace market model. Panel C lets the pooled rent slope differ by the quartile of the worker's expected wage predicted from observed characteristics. Panel D interacts the between rent term with a static 2024 measure of small-unit housing supply constraint and, separately, replaces log rent with the residual from projecting it on access, amenities, and supply. The within component of the job-stock measure reflects half-year changes in the employed-youth stock of a municipality and is not interpreted. None of the access, wage, or supply contrasts is statistically distinguishable from zero; the residual-rent between slope is not an estimate of amenity capitalization. Standard errors in parentheses are clustered two-way by stable residence and workplace geography unless stated otherwise. Estimates are conditional associations, not causal effects.

Table A6: Employment-state model and weighting diagnostics

Panel A. Employment states of persons aged 19–34				
	Wage employed (%)	Other employed (%)	Nonemployed (%)	Persons
2021H2	54.3	5.1	40.5	60,966
2022H1	55.7	5.3	39.0	58,011
2022H2	57.4	5.6	37.0	62,023
2023H1	57.2	5.6	37.1	59,865
2023H2	57.7	5.7	36.6	58,238
2024H1	57.9	5.6	36.6	55,674
2024H2	58.1	5.7	36.3	54,679

Panel B. Multinomial logit of employment state, residence attributes				
	Wage employed (SE)	p-value (wage)	Other employed (SE)	p-value (other)
Standardized rent (z)	0.004 (0.023)	0.855	0.043 (0.038)	0.254
Job openings within 30 km (z)	0.030 (0.039)	0.434	−0.068 (0.049)	0.169
Medical index (z)	−0.013 (0.018)	0.454	0.014 (0.035)	0.693
Cultural index (z)	−0.039 (0.019)	0.042	0.132 (0.031)	<0.001
Education index (z)	0.013 (0.024)	0.599	−0.052 (0.032)	0.102
Transport index (z)	0.009 (0.014)	0.519	−0.082 (0.018)	<0.001

Panel C. Average marginal effects (percentage points)			
	Wage employed	Other employed	Nonemployed
Standardized rent (z)	−0.07	0.20	−0.13
Job openings within 30 km (z)	0.85	−0.45	−0.40
Medical index (z)	−0.32	0.12	0.20
Cultural index (z)	−1.25	0.81	0.45
Education index (z)	0.44	−0.31	−0.13
Transport index (z)	0.47	−0.45	−0.03

Table A6: Employment-state model and weighting diagnostics (continued)

Panel D. Model fit	
	Value
Persons	409,456
Parameters	86
McFadden pseudo-R^2	0.149
AUC for wage employment	0.725
Weighted accuracy	0.702
Residence clusters	290

Panel E. Predicted wage-employment probability and weights						
	Persons	P1	Median	P99	ESS	ESS (% of persons)
All persons 19–34: nonemployed	160,732	0.087	0.479	0.776	122,235	76.0
All persons 19–34: other employed	23,862	0.187	0.680	0.786	17,683	74.1
All persons 19–34: wage employed	224,862	0.139	0.691	0.786	172,221	76.6
Estimation sample: analysis weights	224,862	0.139	0.691	0.786	172,221	76.6
Estimation sample: IPW 1st/99th	224,862	0.104	0.659	0.785	135,009	60.0
Estimation sample: IPW 5th/95th	224,862	0.111	0.665	0.785	149,235	66.4
Estimation sample: overlap weights	224,862	0.112	0.657	0.783	155,081	69.0

Table A6: Employment-state model and weighting diagnostics (continued)

Panel F. Covariate balance, standardized mean differences

	Target mean	Analysis weights	IPW 1st/99th	IPW 5th/95th	Overlap weights
school_2	0.203	−0.333	0.000	−0.080	−0.120
edu_3	0.554	0.260	0.000	0.045	0.009
edu_2	0.430	−0.255	0.003	−0.043	−0.012
age_20_24	0.263	−0.210	0.004	−0.034	−0.028
age_30_34	0.344	0.154	0.002	0.031	0.019
age_25_29	0.352	0.097	−0.005	0.023	0.031
school_4	0.028	−0.055	0.002	0.008	0.026
marital_2	0.177	0.051	0.002	0.020	0.054
opening_access_z_30	0.596	0.045	0.004	0.011	−0.020
transport_index_z	0.438	0.034	−0.001	0.007	−0.010
Maximum absolute SMD, all covariates		0.333	0.035	0.080	0.120

Notes: Panel A gives survey-weighted employment states of all persons aged 19–34 in the seven half-years. Panel B reports the multinomial logit of employment state (base: nonemployed) on standardized residence attributes, with demographic, half-year, and province controls; standard errors are clustered by residence municipality. Panel C converts the coefficients to average marginal effects in percentage points. Panel E summarizes the predicted probability of wage employment and the effective sample size (ESS) under each weighting scheme; inverse-probability weights are trimmed at the 1st/99th or 5th/95th percentiles, and overlap weights multiply by one minus the probability. Panel F lists the ten covariates with the largest standardized mean difference (SMD) between wage workers and all persons aged 19–34 under analysis weights; covariate names follow the estimation code, and z denotes standardization. The reweighting balances observed characteristics only and does not recover the potential workplaces of nonworkers.

Table A7: Reproducibility gates, cross-stage identities, limitations, and claim registry

Panel A. Stage gates

	Verdict	QA checks passed	Binding failures	Nonbinding limitations
E0	PASS	13/13	0	0
E1	PASS	194/194	0	0
E2	CONDITIONAL	28/28	0	2
E3	PASS	21/21	0	1
E4	PASS	31/31	0	5
E5	PASS	33/33	0	1

Panel B. Frozen artifacts

	Files	Existence, size, and hash checks passed
E0	8	8
E1	9	9
E2	9	9
E3	9	9
E4	9	9
E5	10	10
Strategy documents	3	3

Table A7: Reproducibility gates, cross-stage identities, limitations, and claim registry (continued)

Panel C. Cross-stage identities					
	Left value	Right value	Statistic	Tolerance	Passed
E1 full-choice slope vs E2 collapsed slope	−1.153	−1.153	2.31e-05	0.0001	yes
E3 within slope vs E5 unadjusted	0.080	0.080	0	1e-10	yes
E3 between slope vs E5 unadjusted	−1.254	−1.254	0	1e-10	yes
E3 capital-workplace difference vs E5 unadjusted	−5.531	−5.531	0	1e-10	yes
Between slope with job access (E4) vs E3	−1.254	−1.251	0.00235	0.05	yes
Capital-workplace difference with job access (E4) vs E3	−5.531	−5.625	0.017	0.1	yes
Between slope under IPW (E5) vs E3	−1.254	−1.265	1.01	[0.5, 1.5]	yes
Capital-workplace difference under IPW (E5) vs E3	−5.531	−6.028	1.09	[0.5, 1.5]	yes
Sufficient-statistic rows, E3 vs E4	335,839	335,839	0	0.0	yes

Table A7: Reproducibility gates, cross-stage identities, limitations, and claim registry (continued)

Panel C. Cross-stage identities (continued)					
	Left value	Right value	Statistic	Tolerance	Passed
Sufficient-statistic rows, E3 vs E5	335,839	335,839	0	0.0	yes
Choice groups, E3 vs E4	1,596	1,596	0	0.0	yes
Choice groups, E3 vs E5	1,596	1,596	0	0.0	yes
Required frozen artifacts present	57	57	0	0.0	yes

Table A7: Reproducibility gates, cross-stage identities, limitations, and claim registry (continued)

Panel D. Limitation flags			
	Active	Evidence	Required response
Common-support retention (E2)	yes	Central common support retains only 19.7% of the limiting market's alternative-semester cells.	Limit external-validity claims for the common-support estimate.
Wild-score precision in both workplace markets (E2)	yes	The residence-block wild-score interval does not exclude zero in both workplace markets.	Center the market claim on capital workplaces and report the null noncapital result.
Within-municipality coefficient not identified	yes	The pooled within-SGG coefficient is 0.103 in the SGG-FE check and the Mundlak within estimate is statistically insignificant.	Do not claim a short-run causal rent effect.
No exogenous housing-cost shock	yes	No defensible instrument or quasi-experimental housing-cost shock is used.	Use association, gradient, and sorting language.
Seven half-years	yes	The national panel contains seven half-year periods from 2021H2 through 2024H2.	Avoid long-run dynamic or pretrend claims.
Only one mechanism family supported	yes	Only one predeclared mechanism family receives direction, precision, and robustness support.	Report all access, wage, and supply null results.
Selection reweighting on observables only	yes	IPW balances observed covariates but cannot recover nonemployed people's latent workplaces or unobserved selection.	Keep E5 as sensitivity analysis or appendix.
Manuscript overlap audit pending	yes	The manuscripts have not yet been written, so table, figure, caption, and prose overlap cannot yet be inspected.	Repeat the separation audit immediately before either submission.

Table A7: Reproducibility gates, cross-stage identities, limitations, and claim registry (continued)

Panel E. Claim registry of this paper

	Type	Status	Wording
S-CL-01	scope	required	The outcome is the residence municipality chosen by employed wage workers aged 19–34, conditional on their observed workplace; the estimates are conditional associations, not causal effects for all young people.
S-CL-02	baseline	allowed	In the full choice set, the conditional log-rent gradient is negative and does not depend on sampled alternatives, random seeds, a single half-year, or the inference method.
S-CL-03	primary	allowed	The negative rent gradient arises mainly from persistent between-municipality rent differences; the within-municipality coefficient is not statistically distinguishable from zero.
S-CL-04	primary	allowed	The between-municipality rent gradient is negative for capital-region residences and is not detected for noncapital residences; the capital-minus-noncapital contrast persists when the commuter cities adjacent to the capital region are removed from every choice set.
S-CL-04B	boundary	required	Among workers employed in the capital region, the large capital-minus-noncapital contrast depends on a few large noncapital cities near the capital boundary and is reported only together with the exclusion results.
S-CL-05	boundary	required	For workers employed outside the capital region, the capital-minus-noncapital contrast is imprecise and its significance depends on the clustering block.
S-CL-06	secondary	allowed	Workers at high-rent workplace centers show a steeper rent gradient and combine cheaper residences with longer commutes.
S-CL-07	boundary	required	Job access, expected-wage quartiles, and static small-unit housing supply do not account for the rent-gradient differences.
S-CL-08	sensitivity	allowed	Reweighting on observed employment selection does not overturn the core spatial gradients.

Table A7: Reproducibility gates, cross-stage identities, limitations, and claim registry (continued)

Panel E. Claim registry of this paper (continued)			
	Type	Status	Wording
S-CL-09	boundary	required	The capital-region contrast keeps its sign on central common rent support, which retains about one-fifth of capital residence cells, so its external validity is limited.
S-CL-10	implication	allowed	The policy implication is a design principle: youth housing, transport, and job policies should be considered together with the structure of capital-region workplace markets.
S-CL-11	boundary	required	The positive slope for noncapital residences among capital-region workers reflects commuting between large noncapital labor markets near the capital boundary and adjacent outer Gyeonggi, not a preference for expensive housing.
S-CL-12	boundary	required	Differences in the rent gradient by age, sex, and education are not statistically distinguishable from zero.
P-01	prohibited claim	prohibited	Higher rent causes young people to avoid a location.
P-02	prohibited claim	prohibited	Housing costs determine whether Korean youth become employed.
P-03	prohibited claim	prohibited	The market-rent index measures each person's rent-to-income burden.
P-04	prohibited claim	prohibited	The residual rent coefficient identifies amenity capitalization.
P-05	prohibited claim	prohibited	The estimates identify welfare gains from rent subsidies or housing construction.

Table A7: Reproducibility gates, cross-stage identities, limitations, and claim registry (continued)

Panel E. Claim registry of this paper (continued)			
	Type	Status	Wording
P-06	prohibited claim	prohibited	Capital-region residents are about 5.5 more sensitive to rent than noncapital residents.
P-07	prohibited claim	prohibited	Rent has no effect on residential choice within municipalities.
P-08	prohibited claim	prohibited	Reweighting corrects for employment selection.
P-09	prohibited claim	prohibited	Commuting mediates the effect of rent on residential choice.
P-10	prohibited claim	prohibited	Residents of noncapital commuter cities prefer expensive housing.

Notes: Panel A summarizes the quality gates of the frozen pipeline stages E0 (inputs) to E5 (employment selection); a CONDITIONAL verdict records nonbinding limitations without any failed check. Panel B counts the frozen artifacts whose existence, size, and SHA-256 hash were verified; the same files were re-hashed when this paper's workspace was created. Panel C lists the numerical identities that tie the stages together, re-computed for this paper from the frozen files. Panel D reproduces the limitation flags and the response each requires in the text. Panel E reproduces the claim registry of this paper, which revises the pipeline registry after the commuter-city diagnostics; the manuscript audit also enforces its required and prohibited phrases.

Table A8: Demographic heterogeneity in the rent gradient

Panel A. One interaction at a time

	Comparison group (SE)	Base group (SE)	Difference (SE)	p-value
Age 19–29 vs age 30–34	−1.198 (0.261)	−1.092 (0.274)	−0.107 (0.088)	0.226
Female vs male	−1.238 (0.243)	−1.069 (0.287)	−0.168 (0.098)	0.086
Four-year college or more vs less	−1.094 (0.286)	−1.291 (0.243)	0.197 (0.154)	0.201

Panel B. Joint models

	Joint (SE)	p-value (joint)	Joint with capital interaction (SE)	p-value (capital)
Reference-group slope	−1.179 (0.253)	$<$0.001	0.463 (0.342)	0.175
Age 19–29 difference	−0.051 (0.068)	0.456	−0.061 (0.070)	0.380
Female difference	−0.185 (0.105)	0.078	−0.143 (0.107)	0.184
College difference	0.210 (0.157)	0.180	0.261 (0.153)	0.089
Capital minus noncapital residences	n.a.	n.a.	−1.988 (0.430)	$<$0.001

Panel C. Joint Wald tests of the three interactions

	Statistic	Degrees of freedom	p-value
Joint model	3.192	3	0.363
Joint model with capital interaction	3.056	3	0.383

Notes: The full-choice data are collapsed to sufficient statistics by half-year, workplace, residence, and demographic cell, which preserves the conditional likelihood because residence attributes do not vary within a cell. Panel A adds one interaction of log rent with a demographic indicator at a time; the base-group slope is the log-rent coefficient for the omitted group. Panel B enters the three interactions jointly, without and with the capital-region residence interaction; in the joint model with the capital interaction the reference-group slope refers to noncapital residences. Panel C reports Wald tests that the three demographic differences are jointly zero. The demographic differences were predeclared as uncertain in sign (P4) and are reported whether or not they are significant. Standard errors in parentheses are clustered two-way by stable residence and workplace geography unless stated otherwise. Estimates are conditional associations, not causal effects.

Table A9: Distance functional form and deposit-inclusive rent

Panel A. Distance functional form

	Pooled rent slope (SE)	Capital slope (SE)	Noncapital slope (SE)	Difference (SE)	p-value of difference
Log distance (baseline)	−1.153 (0.263)	−1.446 (0.275)	0.534 (0.329)	−1.979 (0.428)	<0.001
Distance bins	−1.021 (0.222)	−1.241 (0.230)	0.526 (0.351)	−1.767 (0.414)	<0.001
Quadratic in log distance	−1.125 (0.248)	−1.364 (0.257)	0.498 (0.355)	−1.862 (0.438)	<0.001

Panel B. Deposit-inclusive rent index

	Deposit-inclusive (SE)	p-value	Monthly-rent index (SE)
Pooled log-rent slope	−1.045 (0.236)	<0.001	−1.153 (0.263)
Capital residence slope	−1.307 (0.247)	<0.001	−1.446 (0.275)
Noncapital residence slope	0.281 (0.261)	0.281	0.534 (0.329)
Capital minus noncapital	−1.588 (0.337)	<0.001	−1.979 (0.428)
Capital-workplace difference, market model	−2.290 (0.552)	<0.001	−4.276 (0.944)
Noncapital-workplace difference, market model	−0.461 (0.243)	0.058	−0.719 (0.450)

Notes: Both panels use the exact sufficient statistics of the full choice set and the conditional logit with residence-province fixed effects. Panel A replaces log distance with distance bins (reference 0–10 km; the same-municipality indicator is kept) or adds a quadratic term. Panel B replaces the market-rate index, built from monthly-rent contracts, with an index that also admits deposit-only (jeonse) leases converted at the same market rate; the choice sets are unchanged, and the last two rows come from the residence-by-workplace market model with pooled rent slopes. Standard errors in parentheses are clustered two-way by stable residence and workplace geography unless stated otherwise. Estimates are conditional associations, not causal effects.

Appendix: Additional Figures

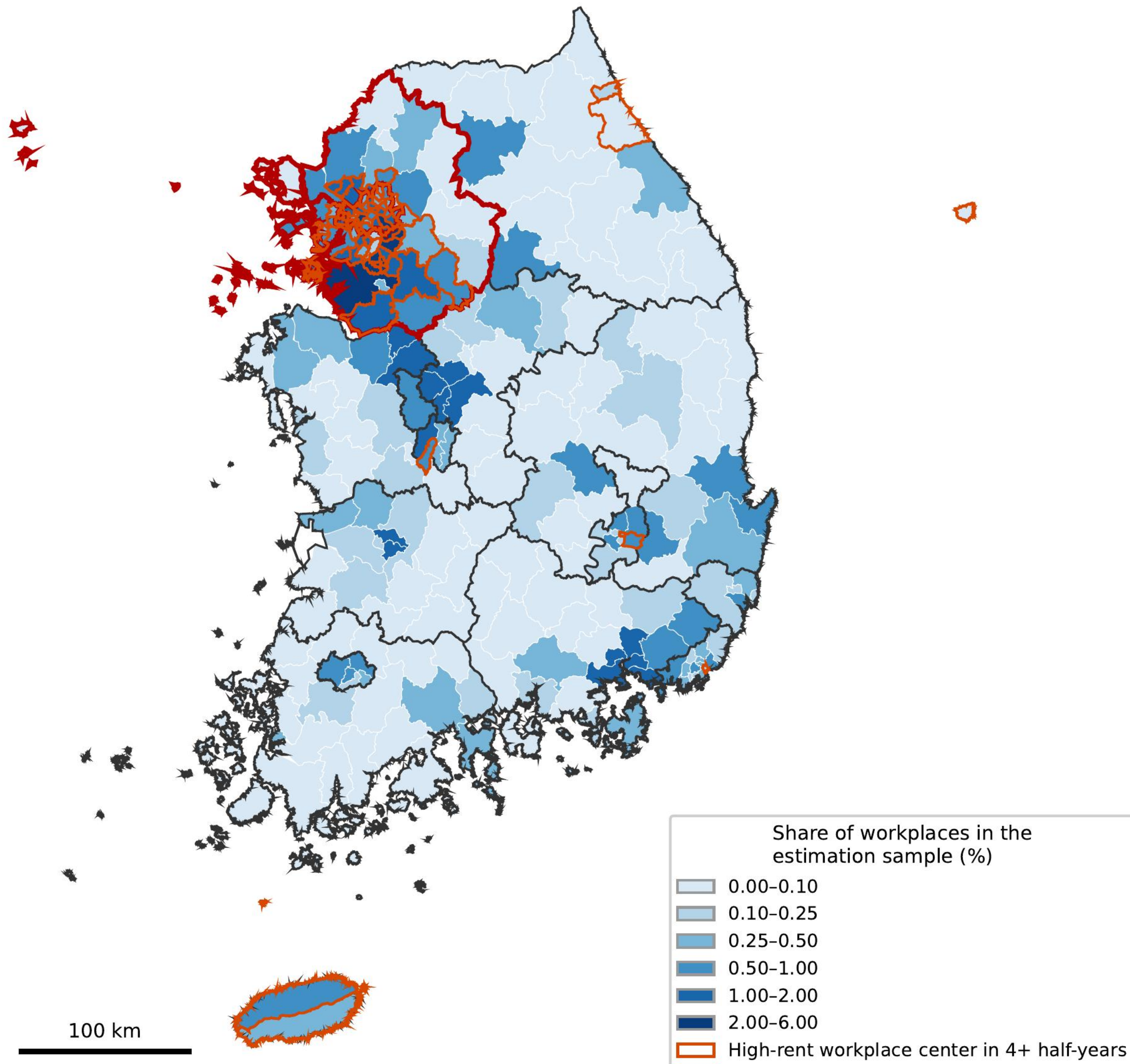

Figure A1: Workplace distribution of the estimation sample and high-rent workplace centers

Notes: Each municipality is shaded by its share of the survey-weighted workplaces of the estimation sample (wage workers aged 19–34 with a rent-eligible residence, 2021H2–2024H2); white municipalities have no sampled workplace. Outlined municipalities are high-rent workplace centers, defined within each half-year as the workplaces in the top quartile of standardized rent (Section 3.4), in at least four of the seven half-years.

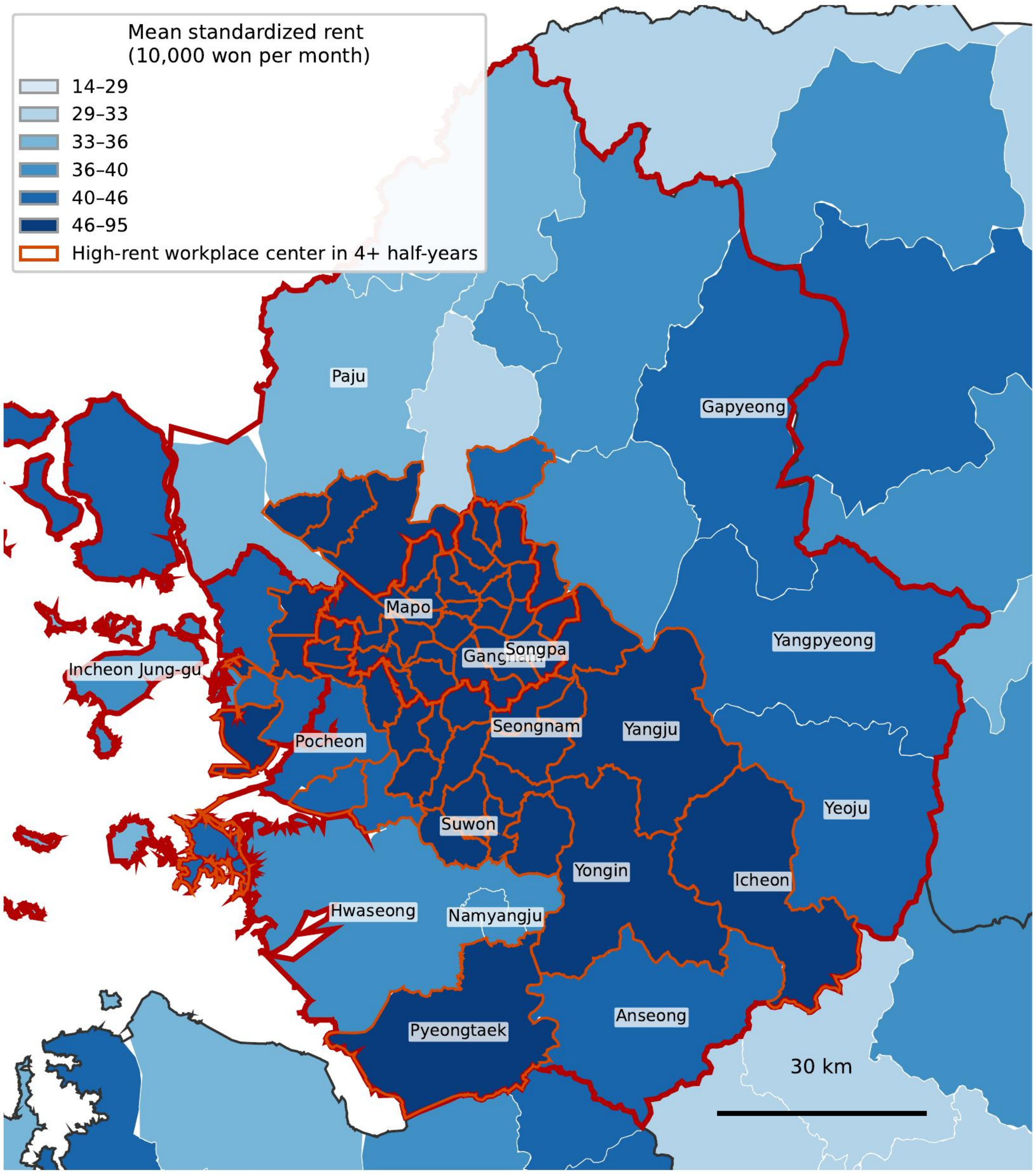

Figure A2: Mean standardized rent within the capital region

Notes: The map enlarges the capital region and its surroundings with the same rent classes as Figure 1, panel (b): the mean of each municipality's standardized 40 m^2 monthly-equivalent rent over its rent-eligible half-years, in 10,000 won per month at 2024 prices. Labels name the municipalities discussed in the text. Outlined municipalities are high-rent workplace centers in at least four of the seven half-years (Appendix Figure A1).

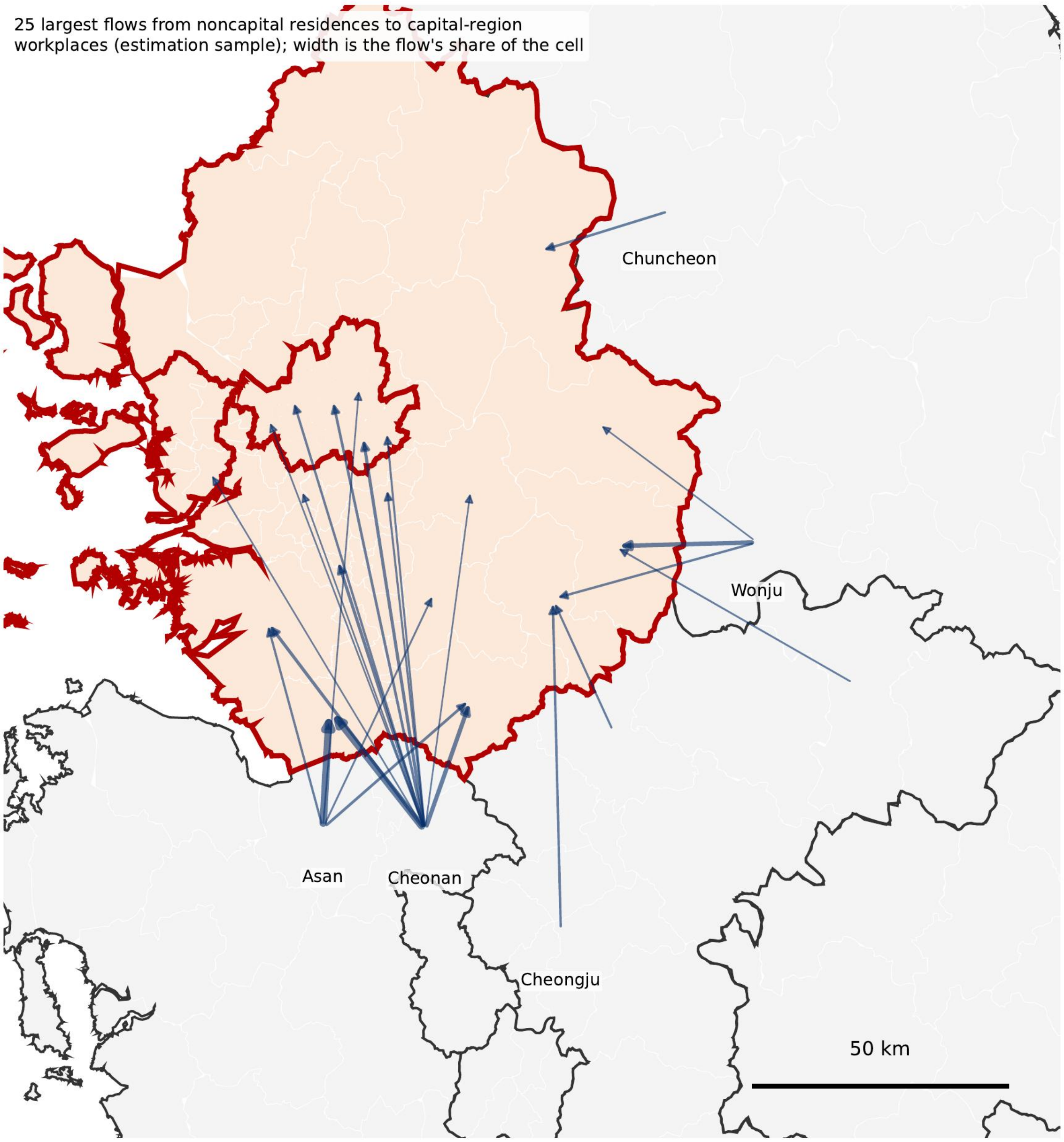

Figure A3: Largest commuting flows from noncapital residences to capital-region workplaces

Notes: Arrows connect the residence and workplace municipalities of the 25 largest flows, by survey weight, among noncapital residents employed in the capital region in the estimation sample; the arrow width is proportional to the flow's share of that market cell. The labeled cities are the five commuter cities of Section 5.3; Seoul is drawn as one unit.