2008
OKSWP0806

Economics Working Paper Series
Department of Economics

OKLAHOMA STATE UNIVERSITY
http://spears.okstate.edu/ecls/

Agglomeration Spillovers and Wage and Housing Cost Gradients across the Urban Hierarchy

Mark D. Partridge
Ohio State University

Dan Rickman
Oklahoma State University

Kamar Ali
University of Saskatchewan

M. Rose Olfert
University of Saskatchewan

Department of Economics
Oklahoma State
University
Stillwater, Oklahoma

339 BUS, Stillwater, OK 74078, Ph 405-744-5110, Fax 405-744-5180

Agglomeration Spillovers and Wage and Housing Cost Gradients across the Urban Hierarchy[*]

by

Mark D. Partridge[1]
Dan S. Rickman[2]
Kamar Ali[3]
M. Rose Olfert[4]

January 29, 2009

Abstract: The tyranny of distance in terms of its effect on median earnings and housing costs is examined for rural and urban U.S. counties. First, we develop a series of distance metrics for an area's remoteness from multiple tiers of the urban hierarchy. Second, we consider geographical access of buyers and sellers through market-potential measures typical of those used in empirical studies of the New Economic Geography. The results reveal penalties of about 5 to 9% for median earnings and 12 to 17% for housing costs due to remoteness from the combined tiers of the urban hierarchy. Differences in market potential also influence factor prices, but these effects are generally smaller than those produced by urban hierarchy distances. Thus, it appears that empirical tests of New Economic Geography models need to consider sources of agglomeration spillovers beyond aggregate market potential. Visually depicting these results using maps illustrates that urban hierarchy distance penalties dominate in the western U.S., but the influence of market potential and urban hierarchy are about equal in the East.

[*] We thank Jordan Rappaport for generously providing Stata code for the estimation of the GMM models and his related advice on modeling.

1. AED Economics, The Ohio State University, 2120 Fyffe Road, Columbus, OH 43210, USA. Phone: 306-966-4037; Fax: 614-688-3622; Email: partridge.27@osu.edu, webpage: http://aede.osu.edu/programs/Swank/.
2. Department of Economics, Oklahoma State University, Stillwater, OK 74078, USA. Phone: 405-744-1434, Fax: 405-744-5180. Email: dan.rickman@okstate.edu.
3. Department of Agricultural Economics, University of Saskatchewan. 51 Campus Drive, Saskatoon, SK, S7N 5A8, CANADA, E-mail: kamar.ali@usask.ca.
4. Department of Agricultural Economics, University of Saskatchewan. 51 Campus Drive, Saskatoon, SK, S7N 5A8, CANADA, Email: rose.olfert@usask.ca. http://www.crerl.usask.ca.

1. INTRODUCTION

An enduring empirical regularity across space is that wages and land prices tend to decline with greater distances from large urban core areas. This pattern applies whether examining land prices or wages within a given urban area or between the hinterlands and successively larger urban centers. Explanations for spatial differences in wages and land prices generally revolve around location responses of households and firms to spatial variations in productivity and amenities.

In the New Economic Geography (NEG), producers and consumers co-locate to exploit plant-level scale economies and to minimize trade transportation costs. Head and Mayer (2004a) note that empirical efforts to assess the NEG have focused on five propositions. Related to factor-price gradients is the first proposition that greater local market potential raises factor prices (and production). This prediction is supported by numerous studies (Davis and Weinstein 2003; Hanson, 2005; Head and Mayer, 2004b; 2006; Ottaviano and Pinelli, 2006; Redding and Venables, 2004).

Aggregate and undifferentiated market potential, however, may be limited in fully reflecting the nature of agglomeration spillovers because it does not consider the spatial nature of urban concentrations and especially the position of these concentrations in the urban hierarchy. There are compelling reasons to consider factor returns within the rich context of precise geographical location and with respect to the urban hierarchy. In particular, classical Central Place Theory (CPT) predicts that successively higher orders of service (with higher demand thresholds) are first available at successively higher levels (tiers) of center. Thus, the industry structure varies across tiers where centers at each tier perform all the functions of the lower-tiered centers plus new ones for which the demand thresholds are first met at that tier. The resulting hierarchical urban system can affect factor prices in a much more rich and nuanced manner than predicted by simple market potential in NEG models. A more complete study of factor-cost gradients over space then requires a spatial and urban-hierarchy frame of reference.

While the systematic nature of the urban hierarchy is well-trodden ground in terms of determining rank and size patterns, empirical studies of spatial spillovers on factor prices are largely absent (Rosenthal and Strange, 2006). Most research has generally examined localized housing costs and wages with respect to the central business district in a given urban area (Lucas and Rossi-Hansberg, 2002), while other studies examined how city size affects wages (Glaeser and Maré, 2001). Exceptions include Brülhart and

Koenig's (2006) assessment of European core-periphery wage gradients and Eaton and Eckstein's (1997) examination of hierarchical city-size effects on human capital formation, factor returns, and hierarchically-derived international development benefits. Yet, there is an absence of studies examining the gradients of factor prices across the urban hierarchical system.

Agglomeration benefits related to urban scale include access to material and labor inputs by businesses, as well as to cultural, recreational, and consumer services by households, all of which bear a direct relation to the specific tier in the urban hierarchy (Christaller, 1933; Krugman, 1996). Likewise, the size of urbanization economies (Jacobs, 1969) including human capital externalities (Glaeser, 1997), are related to urban size. These spatial externalities imply that productivity and, *ceteris paribus*, wages and land rents are directly related to urban tier. Further, they suggest potentially different rates of attenuation of spillover effects across space depending on the tier of the urban center from which the spillovers arise.

For any area then, part of its productivity (dis)advantage is due to its proximity (remoteness) to higher tiers (or larger cities) within the urban hierarchy, and part is due to its aggregate and undifferentiated market potential. Market potential and distance/proximity are not entirely independent as the size and spacing of centers in the urban hierarchy relate to the purchasing power in market areas. Yet, market potential, measured by population or income (within a broader region), does not explicitly consider the spatial distribution of that population and alone, may be an incomplete measure of agglomeration spillovers. For example, whether a given level of market potential is generated by several smaller cities or one large mega center may be an important consideration. Further, market potential alone predicts smooth factor return gradients while hierarchical effects generate discrete changes within the gradient (Brülhart and Koenig, 2006). Also, in a well-established urban system, diminished transport costs imply that other factors such as knowledge spillovers associated with the specific size of cities may dominate those associated with market access (Glaeser and Kohlhase, 2004).

In this paper we systematically investigate how hierarchical geographic proximity and market potential contribute to differences in household earnings and housing costs. Our key contribution is that we explicitly consider both the role of hierarchical distance and standard (NEG-type) market potential effects. Our findings reveal spatial patterns in earnings and housing costs that are systematically related to remoteness from successively higher tiers of the urban hierarchy, with distance effects for housing costs being greater than those for median earnings. These negative hierarchical distance effects are in addition to the standard market potential effects.

We also find that standard market potential effects, consistent with positive agglomeration spillovers, are on average smaller and less pervasive than the urban hierarchy distance effects. Thus, standard market potential is not sufficient to capture the more specifically delineated effects of remoteness from agglomeration effects. Undifferentiated market potential may be more relevant for explaining factor price gradients in densely populated regions such as in Western Europe than for less uniformly densely populated regions. The results suggest that a country's trade and development patterns may be influenced, not only by market potential, but also by the structure of its urban hierarchy, especially for the highest-order services which depend on knowledge spillovers (Glaeser, 2007).

2. REMOTENESS AND THE SPATIAL DISTRIBUTION OF FACTOR PRICES

Numerous studies have linked the spatial distribution of factor prices (and production) to the distribution of demand or access to markets (Davis and Weinstein 2003; Hanson, 2005; Head and Mayer, 2004b; Redding and Venables, 2004). From NEG, greater market access reduces transport costs to final consumers and intermediate suppliers, which along with economies of scale to the firm, facilitates agglomeration economies and raises factor prices. Yet, concentration of economic activity is equally consistent with explanations other than market access (Rosenthal and Strange, 2001).

For example, urbanization economies of Jacobs (1969) such as knowledge and technological spillovers vary with city size. Cultural urban amenities and specialized producer services also may only be available in the largest of cities (though adverse congestion effects also may be prevalent) (Glaeser, 1997). Brülhart and Koenig (2006) found city-size wage effects associated with capital cities across Europe, while Eaton and Eckstein (1997) discover city-size effects on human capital formation and factor returns. These alternative explanations may dominate standard market potential measures in importance for agglomeration and factor price determination (Dekle and Eaton, 1999; Fingleton, 2006; Glaeser and Kohlhase, 2004). A broader context that allows for the influence of both market potential and alternative explanations is one which explicitly considers the hierarchical nature of the urban system, with successively higher orders of function available in successively higher tiers (Christaller, 1933). Proximity to the *particular* city at a *particular* tier may matter for both firm and household location decisions.

2.1 Equilibrium Representation of Remoteness and Factor Prices

Our empirical approach is informed by the spatial general equilibrium framework of Roback (1982) in which both household and firm location decisions are represented. Location characteristics which enhance firm profitability (market demand and closer proximity to larger centers) have a positive effect on wages and land rents. Factor prices also are affected by household preferences for access to consumer services and urban amenities successively available at higher tiers of center, though there may be offsetting congestion effects. Considering the combined firm and household preferences, where households also view closer proximity to agglomeration centers as being, *on net,* utility-enhancing, closer proximity will unambiguously increase rents. The effect on wages, however, depends on the relative magnitudes of firm (labor demand) and household (labor supply) evaluations of distance effects.[1]

Figure 1 illustrates the potential factor price effects à la Roback of distance from markets and from larger cities (undifferentiated by urban tier, for now). For each location, at a given distance, isoutility curves $V(\cdot)$ are upward sloping because higher land rents need to be offset by higher nominal wages to leave the household equally well-off. Likewise, there exists an isocost curve $C(\cdot)$ that is downward sloping because higher rents require lower wages to leave costs unchanged. Thus, an equilibrium combination of w and r exists for each location.[2]

At some arbitrary distance d_1 to the set of higher urban tiers, the equilibrium wage and rental rates equal w_1 and r_1 at point A, where the corresponding isocost and isoutility curves intersect. Assuming that locations at a greater distance from markets and from higher-tiered cities have lower profit potential, a greater distance d_2 shifts the isocost curve leftward. In the absence of household effects, equilibrium wages and rents decline to w_2 and r_2 at Point B.[3] Allowing for only firm distance effects, increased distance would thus result in successive equilibria in the southwest direction from the initial equilibrium—i.e., both wages and rents decline.

The impact of increased distance on the isoutility curve is less clear, as households more likely experience some offsetting positive effects of distance. If distance exerts a *net* negative influence, say for

---

[1] If households, *on net*, perceive closer proximity (due to congestion or disamenties) as utility-eroding, closer proximity will unambiguously lead to higher wages but the effect on rent is ambiguous.

[2] Ottavanio and Pinelli (2006) similarly show how NEG predictions can be cast within a Roback framework for regions of Finland, but they do not consider other possible explanations for agglomeration.

[3] From NEG notions of agglomeration shadows, to the extent that the firm *benefits* from reduced spatial price competition, the leftward shift would be moderated.

example, because of reduced accessibility to urban amenities, V(·) also shifts leftward (as shown in Figure 1), further reducing rents, but partially offsetting the wage decrease represented by B. At point C, the firm-induced negative wage effects dominate the positive wage effects from the household (supply side) effects, to produce a net reduction in the wage rate. In contrast (not shown in Figure 1), if remoteness were household attractive, such as through avoidance of urban congestion costs, V(·) would shift rightward, further reducing wages, but making the net effect on rents ambiguous.

2.2 Conceptualizing the Role of Distance across the Urban Hierarchy

The interregional net agglomeration effects depend on the tiers (sizes) of urban centers, and the distance between these centers and the interacting regions. The theoretical basis for assessing how distance affects location patterns, and the resulting hierarchy of cities, follows from Central Place Theory (Christaller, 1933). Across the hierarchy of cities, agglomeration economies are expected to be greatest in the highest-tier centers. The lowest tier may offer only the most basic services such as a gasoline station. The next higher urban tier would offer all of the services in the lowest tier plus other higher-ordered services such as big box stores. The third tier would offer all of the services found in the second tier plus additional higher-ordered business services such as accountants, attorneys and software consultants. And so on up the urban hierarchy (e.g., see the review by Mulligan, 1984).

Households and firms in the hinterlands and lower-tiered cities must traverse the distance to the nearest higher-tiered cities to access higher-order services, urban amenities, higher-paying jobs, and lower-cost products (i.e., factors related to agglomeration economies). A non-metro area, situated at some distance from each of (say) 4 higher tiers in the hierarchy will find its equilibrium wages to be in part a reflection of agglomeration spillovers from each of the 4 higher-tiered urban centers. We expect wages (and land rents), reflecting firm agglomeration economies, to be greatest in the top tier of the urban hierarchy, declining across the tiers to the lowest level of the hierarchy. Local (non-metro) factor prices will be influenced by the non-metro area's proximity to the nearest urban center through spillovers such as commuting opportunities as well as input-output linkages. If the nearest center is already at the highest tier of the urban hierarchy, we expect no additional influences from other urban centers.

However if the nearest urban center to a given non-metro location $\boldsymbol{i}$ is a tier 3 center, for example, agglomeration economies consistent with a tier 3 center will influence factor prices in the non-metro location *and* there will be an added influence resulting from the non-metro area's remoteness from the tier 4 center. The remoteness effect of the tier 4 center reflects only access to the *incremental* services found at tier 4. For instance, if a business located in location ***i*** needs a patent attorney (assumed to be available at tier 4 or higher), the business need *not* also travel to a tier 4 city to purchase gasoline or basic office supplies (available in the *closer* tier 3 city). They would travel to a tier 4 city only for services that are unavailable in a (closer) tier 3 city. Thus, to capture the influence on factor prices of marginal spillovers from agglomeration economies uniquely present in the tier 4 center, only the *incremental* distance to the tier 4 center should be included. Finally, since each incrementally higher-tier city need to be accessed for only the subset of services that are *not* available in closer lower-tiered cities, Central Place Theory generally implies that marginal distance costs decline when moving up the urban hierarchy.

The distance penalties in terms of lower factor prices for community i in tier j of n tiers of cities in the urban hierarchy can be formally expressed as follows. Each consecutive higher-tier urban area (t, t+1, t+2…n) has successively higher orders of services, more urban amenities, additional agglomerating externalities and higher factor prices. Consistent with a marginal cost measure, the incremental ‘penalty’ then is related to the *difference* in the distances to place t+1 and place t, both measured from place i. For each i, beyond the own level j of urban center ($j<t\leq n$), let the incremental distance to the nearest higher-tier t place equal d^t and the marginal factor price “penalty” associated with greater incremental distance from a tier t place equals φ^t. The different values of φ^t across tiers show that the total distance response has multifaceted segments (not just one linear response across the urban hierarchy).

The total factor price penalties for a given community i in the jth tier can then be depicted as:

(1) $Penalty_{ij} = \sum_t d^t \varphi^t$,

where the summation is over $t = j+1$ to the n^{th} tier. The penalty term reflects the total cumulative disadvantage in terms of lower factor prices due to distance measured across all urban tiers.

Figure 2 illustrates the distance penalty for a location i that is assumed to be at the lowest tier (tier 0), in a 4 tier urban hierarchy. Location i is situated distance d_1 from the next higher level (tier 1) center; distance d_2 from a tier 2 center (incremental distance $d_2 - d_1$); distance d_3 from tier 3 (incremental distance $d_3 - d_2$); and distance d_4 from tier 4 (incremental distance $d_4 - d_3$). The φ terms show the respective marginal factor price penalties. Consistent with Central Place Theory, the figure shows the marginal penalties declining with successively higher urban tiers (though this is empirically verified). P_4 represents the cumulative (reduced) factor price penalty for remoteness from the urban system.

Figure 2 reflects the nonlinear and discrete nature of distance effects in the urban hierarchy and the intervening effects of more proximate higher-tiered urban areas that are below the highest tier. Specifically, in addition to the level of agglomeration economies (and factor price effects) at a particular tier t, remoteness from an even higher tier would carry an added penalty, but that additional disadvantage is only *incrementally* greater than penalties already incurred at closer locations.

2.3 Empirically Defining the Urban Hierarchy

We empirically operationalize distances across the urban hierarchy as follows. For a rural county not part of an urban center, we include first the distance from the population-weighted county center to the population-weighted center of the nearest urban center (defined as a metropolitan area (MA) or micropolitan area (MICRO)). For a county that is part of a MA or part of a MICRO, this first distance is measured from the population-weighted center of the county to the population-weighted center of its *own urban* area.[4] Because MAs and MICROs are often composed of multiple counties defined by a 25% commuting threshold into the urban core, MAs and MICROs can include multiple suburban counties (U.S. Office of Management and Budget, 2000), e.g., metro Atlanta at the extreme contained 29 counties in 2003. *Within* an urban area, the distance effects reflect a host of *intrametropolitan* considerations that affect local factor prices, including congestion, local land use, and zoning. Given our interest in *inter*-urban area effects, we do not focus on this *within*-urban area variable when considering distance penalties.

Beyond the nearest/own urban center of any size, we include the incremental distances to more populous higher-tiered urban centers as described above. First, we include the incremental distance from the county to reach an MA.[5] We also include variables that measure the incremental distance segments to reach an urban center of at least 250K, at least 500K and at least 1.5 million.[6] The largest urban tier generally

---

[4] A micropolitan area is roughly defined as the county(ies) that contain a city of between 10,000-50,000 and other counties with tight commuting links. A MA is similarly defined for counties that surround a city of at least 50,000. We use the 2003 metro/micropolitan area definitions to allow us to use MICROs (first defined in 2003) and to include counties in the metro sample if they had emerging commuting linkages. We desire a spatially broad definition of MAs to isolate *within* MA effects due to commuting patterns versus *inter* urban center interactions due other factors—with inter-MA effects being our focus. See the Census Bureau MA and micropolitan definitions at www.census.gov/population/www/estimates/metrodef.html.

[5] For example, for a county already located in a MA, the incremental value to reach an MA (of any size) would be zero because it already is a MA county. Population-weighted county centroids are from the U.S. Census Bureau. The MA population categories use 1990 population.

[6] If the county is already nearest to a MA that is either larger than or equal to its own size category, then the incremental value is zero. For example, if the county's nearest urban center of any size is already over 250,000

reflects national and primary regional centers, the 500K-1.5m category reflects sub-regional tiers, and the smaller-sizes are lower-tiered centers. We also consider other nonlinearities in the empirical analysis beyond those introduced by allowing distance effects to differ across urban tiers.

For any location the *sum* of the (incremental) distances equals the distance to a MA of at least 1.5 million people; we are measuring, in segments, the cost of reaching the nearest *highest*-tiered city. Figure 2 shows that the *effect* of this distance is disaggregated by segments to each intervening tier city, yielding an overall non-linear relationship, i.e., distance effects vary depending on which urban tier is being considered.

Figure 3 illustrates the distance measurements for Carbon County, a rural Wyoming county containing about 8,500 residents. It's nearest urban area is Albany County (a MICRO), containing the city of Laramie, located 129kms away. The nearest MA is Casper, located 143kms away, an incremental distance of 14kms (143-129). Because Casper is a small MA, with about 67 thousand people, Carbon County residents and businesses will also be influenced by agglomeration economies in the nearest larger MA of at least 250K, which is Ft. Collins, Colorado, with population just at 250 thousand. Ft. Collins is 210kms from Carbon County, but the incremental distance to agglomeration economies not present in Casper is 67kms (210-143). The nearest MA of at least 500K, the next higher tier above Ft. Collins, is Denver, Colorado, with roughly 2.5 million people. Denver is 282kms away from Carbon County, an incremental distance of 72kms (282-210). Because Denver has over 1.5 million people, it is considered in the top urban tier and there are no added factor price distance penalties (incremental distance to reach a MA of at least 1.5 million equals zero). Yet in sensitivity analysis, we also assess whether distances to the three largest U.S. MAs matter (Chicago, Los Angeles, and New York).

Corresponding to Figure 2, the following distances apply to Carbon County: d_1 = 129 (Laramie-Rawlins), d_2= 14 (Casper-Laramie), d_3= 67 (Fort Collins-Casper) and d_4 = 72 (Denver-Fort Collins). Reflecting the urban hierarchy, factor prices in Carbon County are 'discounted' relative to those in Laramie as a result of the distance d_1 in Figure 2 that must be traversed to access Laramie's agglomeration benefits. Likewise, they are further discounted until the highest urban tier (Denver) is reached.

2.4 Aggregate Market Potential

A long-standing representation of market access is the concept of market potential introduced by

---

people and 40kms away, then the nearest urban center is 40kms away and the two incremental distance values for nearest MA of any size and the nearest MA> 250,000 are both equal to zero.

Harris (1954), typically measured by a proxy for total market demand inversely weighted by distance. Not being explicitly derived from any theoretical model, market potential is interpreted as a reduced-form measure of market access to suppliers and customers (Head and Mayer, 2006). An alternative measure of market access is explicitly derived from NEG, which accounts for *real* price effects of potential competitors (Head and Mayer, 2004a; Hanson, 2005). Yet, atheoretical measures of market demand are commonly used and appear to perform well compared to the explicitly derived measures that account for NEG real price effects. For U.S. counties, Knapp (2006) finds a correlation of 0.95 between a 'Harris' measure of market potential and a NEG measure of supplier access which accounts for real regional price differences. This is not surprising for the U.S. because its internal openness would be expected to lead to less regional price variation in traded goods, resulting in small *real* price differentials in the underlying theoretical measure.[7] Given the close similarities, we use a market potential variable akin to Harris.

Although the effects of market potential and proximity in the urban system are related, they can yield different theoretical predictions. Consider the following two scenarios where we assume the rural populations are identical. In the first, there are two cities of 250K each, one located 100kms to the north and the other 100.1kms to the south (and no other nearby urban centers). In the second scenario, there is one city of 500K located 100kms to the north (and no other nearby urban center). All else equal, the market potential is virtually the same in both scenarios assuming similar income across these cities. However, in our urban system approach, the first scenario leads to the nearest urban center of at least 250K being 100kms away and it is farther yet to reach an urban center of at least 500K in which the larger center has higher-order services and infrastructure (e.g., airports). In the second scenario, it is only 100kms to reach an urban area of at least 500K residents, which would give the county a different factor price effect than in the first scenario in the urban system approach.

In sum, when we include both the hierarchy-specific distances and market potential variables, we are assessing the extent to which the standard undifferentiated market potential measures alone are adequate in capturing the urban agglomeration spillover effects on factor prices. Factor prices should smoothly

---

[7] When examining European data, Head and Mayer (2006) find that the Harris measure of market potential produced results similar to the theoretical measure of real market potential, while Head and Mayer (2004b) find the measure of real market potential to underperform the Harris measure. The Harris measure of market potential might also be thought of as an instrument for real market potential of NEG, but should be cautiously interpreted (Head and Mayer, 2004a). Hanson (2005) also uses a Harris measure, though he adds other variables to reflect NEG features.

diminish as distance reduces market potential. But *additional* urban hierarchy effects should appear as discrete shifts in the factor price gradients depending on an area's proximity to varying urban tiers.

2.5 Descriptive Statistics: The Urban Hierarchy, Market Potential, and Factor Prices

To illustrate how distance affects factor returns, Figures 4a and 4b show the simple relationships between U.S. nonmetropolitan county household earnings and housing costs, and distances from each tier of the urban hierarchy. Our measures of earnings and housing costs are from Census 2000 and are described in the next section. The figures show that nonmetro earnings and housing costs generally consistently decline with distance from all higher tiers of urban areas. The gradients appear steeper at shorter distances, flattening out at greater distances, in which the gradient for nearest urban area ends at a shorter distance than the others. Yet, over significant ranges of distance the gradients are approximately linear. A similar price pattern emerges for smaller MAs when considering distance to higher-tiered urban areas (not shown).

These declines also could be interpreted as consistent with the factor price declines found to be associated with declining market potential (Hanson, 2005). Thus, to assess whether we can empirically distinguish urban hierarchical distance effects from market potential as suggested by our discussion in Section 2.4, Table 1 reports correlation coefficients between our urban hierarchy distances and measures of market potential. As fully described in Section 3, following Hanson (2005), we measure market potential by total personal income within surrounding concentric rings at varying distances from the county.

The correlation coefficients in Table 1 reveal the expected negative relationships between distances from higher-tiered urban areas and the market potential measures. The low absolute values indicate that these are distinct measures. Even when examining the sub-samples of counties used in the empirical analysis, the largest negative correlation is 0.32 (not shown). Though these results support the importance of urban hierarchy, there could be multiple intervening factors underlying the results in Figures 4a and 4b, for which we now turn to regression analysis.

3. ECONOMETRIC IMPLEMENTATION

Our sample consists of over 3,000 U.S. counties (excluding those in Alaska and Hawaii), which we separate into four sub-samples. The first sub-sample contains nonmetropolitan counties (i.e., not part of a MA), including sparsely populated rural areas, towns, and small cities of less than 50,000. In a closely related second sub-sample, we remove counties that belong to micropolitan areas (MICRO) from the nonmetropolitan sample to form a "rural" subsample. Removing counties with tight links to the small cities (10,000-50,000) that comprise MICROs allows us to assess whether the nonmetro sample accurately represents rural areas. The final two sub-samples are those that are parts of MAs areas with less than 250K ("small" MAs) and those that are parts of MAs of more than 250K ("medium" and "large" MAs).[8, 9] The 250 thousand population threshold is selected because it creates two approximately equal sized metro samples—though in sensitivity analysis using 500 thousand as the threshold yields almost identical conclusions. We refer to both MAs and MICROs as urban centers in our discussion.

We expect different equilibrium relationships between remoteness and factor prices across these four sub-samples, such that pooling them would create complicated nonlinearities and heterogeneities. For instance, compared to a small urban center, a rural county's wages and land rents would be affected differently by proximity to its medium and high-tiered urban centers because rural areas have a different industry composition than urban counties (e.g., Desmet and Rossi-Hansberg, 2007). Places in each subsample likely also have different transport technologies, commuting patterns, and congestion levels.

The natural logs of median 1999 county earnings for employed residents (EARN) and median 2000 housing costs (HCOST) are the primary dependent variables. Both variables are derived from U.S. Census 2000 SF3 files. Although our theory indicates land rents should be examined, housing costs should be a good proxy for land rents because spatial differences in quality-adjusted residential housing prices result primarily from differences in the included land values (Davis and Palumbo, 2006).

The variable HCOST is defined as the log of weighted average median gross rent ($ per month) of owner and renter-occupied housing units for 2000 (Gabriel et al., 2003). For owner-occupied units, median housing prices are converted into imputed annual rent using a discount rate of 7.85% (Peiser and Smith, 1985). The monthly average of this amount along with the median monthly rent for the renter-occupied units, weighted by the shares of owner- and renter-occupied houses, is our median housing cost variable. Using median values implies that both dependent variables capture the typical worker/resident's

---

[8] Counties have the advantage that their borders are not affected by recent growth experiences (such as MAs). We follow the U.S. Bureau of Economic Analysis in merging independent cities with the surrounding county to form a more functional region (mostly in Virginia). Forty three mostly small rural counties are omitted due to the lack of economic data. See Partridge and Rickman (2006) for sample details.

[9] We use aggregate county-level data because individual census level housing and earnings data do not contain geo-identifiers for the two-thirds of counties that are nonmetropolitan, which would preclude analysis of them (i.e., we would be forced to only consider metropolitan areas). Examples of studies that use individual worker and housing data include Bloomquist et al. (1988) and Gabriel and Rosenthal (2004). Studies like ours that utilize aggregate-level data to capture finer spatial detail include Head and Mayer (2006) and Hanson (2005).

earnings and housing cost, consistent with the representative individual approach in Section 2.[10]

Following the quality-of-life literature (Gyourko et al., 1999) and some empirical NEG studies (Redding and Venables, 2004), we assume that the spatial earnings and housing cost relationships hold at all periods of time. This study differs from typical hedonic quality-of-life studies in that spatial proximity rather than natural amenities is our primary focus. This spatial emphasis could create two complications. First, at the county level, some of the workforce and housing "quality" variables may be correlated with urban distance and market potential, creating multicollinearity. Second, endogeneity bias could be introduced if some of the explanatory variables are correlated with the residual, which could occur if there are omitted factors that influence earnings and/or housing costs and these omitted factors are also correlated with our local measures of the workforce and/or the housing stock.

The empirical approach consists of examining alternative specifications to assess robustness and econometric concerns. We begin with a parsimonious model that includes only measures of distance, fixed-location measures such as climate, deep lags of the county's own and regional population, and state fixed effects. This implicit reduced-form model mitigates multicollinearity and alleviates endogeneity concerns because these explanatory variables are exogenous or predetermined. Subsequently, we add market potential variables, forming the base model. Numerous other models are then estimated to assess the sensitivity of the base model results. Partially previewing our results, the key distance and market potential conclusions are not altered by the extensive sensitivity analysis, which should greatly alleviate any concerns that econometric misspecification underlies the results.

We specify hedonic equations for the log of median earnings (EARN) and the log of housing costs (HCOST). Except for the distance measures, most of the explanatory variables are from the U.S. Census Bureau SF3 file (details in Appendix Table 1). In each of the four sub-samples, for county *i*, located in state *s*, the base parsimonious county earnings and housing cost equations are:

(2) $EARN_{ist} = \alpha^{W} + \varphi^{W} GEOG_{ist\text{-}l} + \gamma^{W} \mathrm{AMENITY}_{\mathrm{is}} + \sigma^{W}{}_{s} + \varepsilon^{W}{}_{ist},$

(3) $HCOST_{ist} = \alpha^{H} + \varphi^{H} GEOG_{ist\text{-}l} + \gamma^{H} \mathrm{AMENITY}_{\mathrm{is}} + \sigma^{H}{}_{s} + \varepsilon^{H}{}_{ist}.$

The **GEOG** vector contains distance measures to different tiers in the urban hierarchy (see Section 2.3),

---

[10] Median earnings are the preferred measure over other possible measures. For instance, per-capita income would not exactly capture the notion of wages because it includes capital income. Likewise, besides not representing the median *worker*, the average wage per job in the county could be greatly skewed, especially in rural areas with seasonal and part-time work. For example, this measure would count a situation of a person working in one full-time job and two infrequent part-time jobs as three unique jobs that would likely average a relatively low wage (and because it is place of work, it would not reflect commuting opportunities that affect the median worker).

lagged (NEG) market potential variables, and other lagged variables reflecting the scale of the county's region; the **AMENITY** vector contains measures of natural amenities; and state fixed effects σ_s account for common factors within a state. A key advantage of these specifications is the exogenous/predetermined nature of the explanatory variables. The regression coefficients are α, φ, and γ; and ε is the residual.

In addition to the distance and market potential variables described in Section 2, the **GEOG** vector includes the county's population. If the county is part of an urban area, we include the total urban-area population (MICRO or MA), we use population of the *nearest* urban area in the nonmetro samples. In sensitivity analysis, we include the incremental population of higher-ordered urban centers beyond the nearest urban area. County size measured in square miles is also included. Greater county size suggests lower employment density and more space to construct housing. It also implies greater distance *within* the county to reach higher-order services and customers that may affect *within*-region agglomeration spillovers.

Market potential is represented by aggregate household income in surrounding 100-200, 200-300, 300-400, and 400-500km rings measured from the population-weighted center of the county, which allows for a flexible functional form.[11] We use lagged (1989) measures of market potential to mitigate any direct endogeneity.[12] We do not include aggregate income within 100km of the county of interest because this would introduce endogeneity by definition as the largest component of household income is earnings, one of our dependent variables. Many NEG studies similarly remove the own region or the own country in calculating market potential to mitigate this concern (Redding and Venables, 2004; Head and Mayer, 2006; Knaap, 2006). In our case, the own county population and population of the own or nearest urban center account for more localized market potential.

To avoid omitted variable bias, we control for potential causes of labor supply shifts and other location attributes that may affect firm location (Fingleton, 2004). Natural **AMENITIES**, reflecting 'first nature' location effects, are measured by four climate variables, a 1 to 24 scale variable related to topography, percent water area, and three indicators representing location within 50kms of the Pacific Ocean, Atlantic Ocean, and Great Lakes. State fixed effects account for state factors such as settlement period, policy

---

[11] A county's household income was included in a ring if its own population-weighted centroid fell in that ring. A 500km limit follows Hanson's (2005, p. 20) conclusion that the effects of shocks to market potential barely extend beyond 400km. Another way would be to inversely weight neighboring county income by the distance from the county, but our approach is more flexible because we do not have to find an optimal weighting scheme.

[12] In addition to possible correlation in measurement error, there is a possibility that there is an omitted factor or shock that contemporaneously increases wages and housing prices in both the county and the broad region, which creates endogeneity bias. Lagging the market potential measure 10 years avoids such direct endogeneity.

differences, or natural resource availability. With state fixed effects included, the other regression coefficients are interpreted as average responses for *within*-state changes in the explanatory variables.

The county residual is assumed to be spatially correlated with neighboring counties in which the strength of the correlation is inversely related to the distance between the two counties. We use a generalized method of moments (GMM) procedure to produce t-statistics that are robust to general forms of cross-sectional spillovers or spatial autocorrelation (Conley, 1999; Rappaport, 2004).[13]

4. EMPIRICAL RESULTS

Descriptive statistics for the full sample are reported in Appendix Table 1, while those for selected sub-samples are reported in Appendix Table 2. Table 2A contains the base EARN regression results for counties located in nonmetropolitan areas, rural areas, small MAs (< 250K population), and large MAs (> 250K), whereas Table 2B contains the corresponding base results for HCOST. In both the household earnings and housing costs models, Chow tests strongly indicate that the nonmetropolitan, small and large MA models should be estimated separately (respectively, F=14.71, p=.0001; F=15.73, p=.0001).

We first describe the results of the parsimonious model that includes only urban hierarchical distances, lagged population, amenity, and state fixed effects. This is followed by a second (the 'base') specification that also includes the market potential variables. Section 4.2 then reports the size of the urban hierarchy distance penalties and market potential effects including an extensive discussion of sensitivity analysis.

4.1 Regression Results

Columns 1-4 in both Tables 2A and 2B show the results from our most parsimonious model. Across the nonmetropolitan, rural, and both MA samples, distance to higher-tiered urban centers is associated with lower earnings and housing costs. Distance significantly affects both median earnings and housing costs as evidenced both by the individual t-statistics for the distance variables, and the F-statistics at the bottom of the table for the joint significance of all of distance coefficients and the joint significance of only the incremental distance variables. The results also are similar across the nonmetropolitan and matching rural sub-sample, suggesting that nonmetro counties reasonably depict the 'hinterlands.'

Regarding the sizes of the coefficients, housing costs are most associated with urban hierarchy distance. For every km farther away a nonmetropolitan county is from its nearest urban center, median earnings fall by about 0.08%, while median housing costs decline about 0.23%. Yet, we find that counties

---

[13] The bandwidth is 200kms, after which we follow convention and assume no correlation in county residuals.

located in both small and medium-sized MAs (with less than 1.5 million 1990 population) are more affected by greater incremental distance from their higher-tiered urban areas than are nonmetropolitan counties.[14,15]

Although our focus is not on the *intra*-urban area distance variables (e.g., distance to the center of own metro variable for a suburban county), the two MA samples indicate the expected strong statistically significant inverse relationship between housing costs and distance to the urban core. Conversely, for median earnings, distance to the own urban core is statistically significant (negative) only in the large MA sample. The own county and nearest/own MA population coefficients suggest that while larger population in the immediate area increases earnings and housing costs in nonmetropolitan counties, these localized effects are more ambiguous in MAs. One tentative conclusion is that in MAs, the positive gains from greater market potential are at least somewhat offset by other effects such as congestion or spatial competition. Median earnings in counties located in nonmetro areas and small MAs are inversely associated with square miles at the 10% level, while medium and large MA housing costs are negatively associated, suggesting heterogeneity in *within region* agglomeration/congestion effects of local distances.

Becoming our base model, columns 5-8 in each table report the results after adding the market potential variables to the parsimonious models (cols. 1-4). This model addresses the key question of whether actual distances to tier-specific urban centers have any factor price effects beyond proximity to aggregate larger customer and input markets, however spatially distributed. Jointly, the incremental distance variable coefficients remain highly significant across all four models, being especially robust in the MA samples, though the magnitude and statistical significance of the coefficients for the nonmetropolitan earnings and housing cost regressions are marginally smaller than in column 1.

In terms of Figure 1, the negative coefficients for both housing costs and earnings indicate that greater distances result in successive equilibria located southwest of the initial equilibrium. That is, either a) both firms and households incur negative distance effects with those of firms being greater than for households, or b) negative distance effects experienced by firms are greater than net positive effects for households.

Consistent with NEG predictions on productivity and profits, the market potential coefficients suggest

---

[14] If the distance terms do not accurately reflect actual travel time, this measurement error would bias the distance coefficients towards zero—working against finding statistically significant distance effects.

[15] Although actual distance does not capture our theoretical notion of a *marginal* or incremental cost to reach a next higher-tiered city, using actual distances in place of incremental distance still results in the distance coefficients being jointly significant, though the individual coefficients are measured with less precision. This is not surprising because these measures have much more multicollinearity—if it is a long distance to reach a MA of at least 500K, it is by definition going to be a long distance to reach a MA of at least 1.5m.

that greater neighboring aggregate household income between 100-200kms and 300-400kms are statistically significant and positively related to nonmetro housing costs and median earnings, with a similar pattern for the rural results. Likewise, own-county population is positive and significant. Surprisingly, for the MA samples, the market potential coefficients are generally negatively related to median earnings and housing costs. Taken together, the aggregate household income variables are jointly significant in the nonmetropolitan and rural samples, but only for the medium/large MA sample are they significant for housing costs. This suggests that the urban hierarchy proximity effects are more consequential in affecting median earnings and housing costs than are the market potential variables.

4.2 Cumulative Urban Hierarchy Distance Penalties and Market Potential Impacts

Table 3 presents the combined distance penalties for median earnings and housing costs, measured at the mean distances (Appendix Table 2), i.e., the cumulative effects of being farther away from higher-tiered urban areas. For each case, we report two distance penalty estimates: 1) a ‘conservative’ measure that uses only distance coefficients that are statistically significant at the 10% level; and 2) a measure that uses all of the distance coefficients regardless of their precision of the individual coefficient estimates. The total urban hierarchy distance penalty is the summation across relevant higher-tiered urban area coefficients.[16] For the two MA samples, the penalty does not include the distance to the own-MA center, since that is an *intra-urban area* effect.[17] We are interested in accessibility to higher-ordered services in the nearest higher-tiered urban areas—not *within* a given urban area (say between the suburbs and the principal city).

Parsimonious and Base Distance Models. Table 3 reports the size and the robustness of the distance penalties for variations of the specification to assess the sensitivity of the overall conclusions. The distance penalties are first reported for our most parsimonious model in Row 1 (shown in cols 1-4 of Tables 2A and 2B). For this case, the typical nonmetropolitan county at the mean distance from its urban tiers has about 7.6% lower median household earnings and 19.8% lower housing costs than an otherwise equal county immediately adjacent to its highest-tier city (regardless of considering only coefficients significant at the

---

[16] The results correspond to taking the coefficients from Table 2A multiplied by the means from Appendix Table 2 and summing the result to obtain a distance effect δ. To account for the log-linear specification, we take 100(exp(δ)-1) to convert the figure into percentages. For example, for the nonmetro median-earnings distance penalty when using all coefficients, the results correspond to taking the sum of –7.7E-04(41.07) – 3.0E-04(55.4) – 2.5E-04(66.8) – 1.4E-04(42.89) – 8.4E-05(89.03) and then taking the exponential of the sum δ to convert to -7.58 percentage points.

[17] We include this impact in the nonmetro sample for MICROs. In MICROs, this has little practical influence because the distance to the nearest urban area is zero in most cases. This follows because the typical MICRO is one county, or distance to its nearest urban area would equal zero. The rural sample provides a good comparison because MICROs are omitted and the distance to the nearest urban area (either a MICRO or a MA) is usually much larger.

10% level or using all coefficients). Moreover, in all cases, the t-statistics fall in the 5 to 9 range, suggesting the distance-penalties are precisely estimated.

A pattern in Row 1 that generally applies for other cases is that the two sets of metropolitan distance penalties for earnings are greater than the corresponding penalties for the rural and nonmetropolitan samples, but the urban distance penalties are smaller in magnitude for housing costs. The relatively large distance penalties for the urban sample suggest that even medium-sized MAs have industry compositions that benefit from proximity to urban services found in subsequently larger metro areas. The distance penalties generally are similar regardless of whether we use all the distance coefficients or only those that are significant at the 10% level.

The **base** model's (including market potential variables) penalties are shown in Row 2, which are derived using results from cols 5-8 of Tables 2A-2B. The distance penalties are modestly reduced in the rural and nonmetro samples, but they are approximately the same in the two MA samples. Specifically, the base results suggest a distance penalty for rural and nonmetropolitan areas in the 5% to 8% range, while the penalty is estimated to be 13% to 17% for housing costs. For MAs, the earnings distance penalty is estimated to fall in the 8% to 9% range, while it is in the 12% to 15% range for housing costs. In sum, these results illustrate that proximity to the county's specific higher-tiered urban areas has strong effects that are distinct from the effects of simple market potential, suggesting that tests of NEG models that focus only on market potential are missing other forms of agglomeration spillovers, some of which may even be consistent with elements of the NEG (i.e., NEG market potential models may understate spatial spillovers).

Nonlinearities in Distance Segments. As our first appraisal of the sensitivity of the results, Row 3 adds to the base model (Row 2) quadratic terms for all of the distance variables to assess whether allowing for added (within-segment) nonlinearities changes our base conclusions. We further consider potential nonlinearities by then adding cubic terms for all of the distance terms to the model described for Row 3 to produce the results in Row 4. The square and cubic terms are jointly significant at the 1% level. The cubic model's adjusted R^2 statistics always exceeds the R^2 for the quadratic model, which in turn exceeds that from the base model (not shown). For both median earnings and housing costs, the distance penalties (Table 3) are larger when considering these nonlinearities (especially the cubic results), though the increase is less notable for the large MA sample. Thus, our base results appear to understate the distance penalties. As shown in Table 4, the corresponding market potential effects are not affected (described below).

As expected, the nonlinear models suggest distance has a declining marginal effect. For example, when using all the coefficients, the nonmetropolitan distance penalty using the cubic model is 18.7% when examining distances measured at *twice* their mean level versus 10.8% when considering distances measured *at* their mean values, while the corresponding values for the median housing costs are 44.3% and 30.1%. While our main conclusions are unaffected, there appears to be declining marginal transportation costs in less congested or more remote environments.[18]

Conversely, many empirical international trade studies use a power distance function—log distance—rather than an exponential function (or linear distance) to measure distance costs (Head and Mayer, 2004a). Yet, there are theoretical problems with the power function when there are relative low transportation costs and short distances, as in our case.[19] Nonetheless, Row 5 investigates the effect of replacing the linear distances with their log values. First, unlike the case where the square and cubic models improved the adjusted R^2 over the base model, the log distance specification led to smaller R^2 values in 6 of the 8 base EARN and HCOST specifications (not shown). Compared to the base model, the distance penalty estimates are smaller using log distance, which is not surprising given that it produced an inferior fit. An exception is the rural earnings distance penalty is about 50% larger using the log distance, which is one of the cases where the adjusted R^2 improved over the base specification.[20]

Sorting and related effects. One of the reasons for higher urban factor returns, especially earnings, is sorting of higher-skilled workers into urban areas. Combes et al. (forthcoming) find that sorting is an important explanation for wage agglomeration effects using French data, though Glaeser and Maré (2001) find more modest sorting effects using U.S. data. One of the likely causes for the sorting of higher-ability

---

[18] Though our general pattern of results are not affected using the cubic model, it would require reporting 15 distance coefficients (for the linear and nonlinear terms) per model. We choose not to use it as our base model to keep the volume of reported results manageable.

[19] Theoretically supporting the choice of the exponential distance function, Sen and Smith (1995; p. 118) find that the realized spatial interactions are almost always accurately approximated by an exponential distance function (linear distances when the dependent variable is in log form). For example, the power function implies that wage *levels* (not log wages) are related to distance d through the following function $d^{-\theta}$ ($\theta>0$). However, as $d \rightarrow 0$, $d^{-\theta} \rightarrow \infty$. Thus, even when there is small distance costs θ, there will still be 'overwhelmingly' greater attraction at shorter distances than at greater distances (Sen and Smith, 1995; p. 94). In our model, this is quite problematic because counties are regularly neighbors and transportation costs are relatively small. Conversely, in international trade, relative transportation (and transactions) costs can be much larger and distances are almost always much greater.

[20] Following a 'nonparametric' approach used by Eaton and Kortum (2002), we also used 5 different categorical distance groupings for each urban tier for a grand total of 25 different distance categories (not shown). In this case, the estimated distance penalties were comparable to what is reported for the quadratic and cubic results. Yet, the resulting adjusted R^2 statistics tended to be less than the quadratic and cubic cases. The categorical distance results further suggest that if anything, our base model's distance penalty estimates are on the low side.

workers into larger urban areas is underlying agglomeration effects of thicker labor markets which improve firm-worker matches, as well as provide better consumption opportunities.

By controlling for various education and demographic attributes, we can assess the role of sorting in the impact of urban access on earnings' distance penalties. For similar reasons, we also add housing attributes in the housing cost specification. As already noted, caution should be exercised when including the labor and housing quality measures because their inclusion may increase the potential for endogeneity bias and multicollinearity. For this reason we consider these measures *after* first estimating the parsimonious models that should be free of these econometric concerns. In assessing the role of sorting, we estimate models that employ deep lags of the "quality" measures to obtain predetermined variables to eliminate or mitigate potential endogeneity. We compare the results using 2000 values of the quality measures. Robustness across the parsimonious, and the fuller 1990 and 2000 specifications, would be compelling evidence that our key conclusions are not driven by omitted variables, endogeneity bias, or multicollinearity.

The full model estimates are reported in Row 6 using the attributes measured in 1990 and in Row 9 using 2000 measures (see the notes to Tables 3 and 4 and Appendix Table 1 for variable definitions).[21] For both the 1990 and 2000 full models, the nonmetro and rural results are almost entirely unaffected by adding these variables (with the very minor exception of nonmetro housing cost results), suggesting that sorting behavior does not explain their distance penalties. Conversely, small MA and large MA distance penalties are about halved when including these attributes. The results suggest that factor prices in MAs are affected by access to higher-ordered urban areas through sorting of higher-ability workers. For example, higher-ability workers may be more likely to locate in smaller and medium-size MAs if there is closer access to (even) larger urban areas. Further research should examine this issue more closely, especially in other countries with differing mobility rates.

Other Specification Changes. We next estimate additional specifications to assess whether the results are sensitive to potential factors such as multicollinearity. First, Row 7 removes all of the market potential variables from the model reported in Row 6, while Row 8 removes all of the distance terms from the model used in Row 6. Rows 10-11 report the corresponding models shown in Rows 7-8, using instead 2000 values

---

[21] To control for labor quality, we include either 1990 or 2000 average county-level measures of five racial composition variables, percent married, six age distribution variables, percent female, percent with disabilities, and four education variables (see Appendix Table 1 for details). The average county-level housing quality measures include the 1990 (or 2000) median number of rooms, age of the structure, the ratio of bedrooms to total rooms, the share housing units that are mobile homes, and the shares with complete plumbing and kitchen.

of the workforce and housing quality variables. Compared to the fully specified models, omitting the market potential variables (rows 6 and 9) has a very minor influence on the distance penalty estimates in Rows 7 and 10 (penalty estimates show very modest increases for rural and nonmetro housing costs). These results again indicate that actual geographical position of an area with respect to its urban hierarchy has an influence on factor prices independent of market potential.

The distance penalties also may be understated if there is an even higher urban tier above the current 1.5 million resident tier. To consider this possibility, Row 12 reports the results when we add incremental distances to Chicago, Los Angeles, and New York (assuming they represent an even higher tier) to the base model.[22] The corresponding regression coefficients on the "big-3" distance variables were significant at the 10% level or better in all cases except the nonmetro housing cost model (not shown).[23] The resulting distance penalties reported in Row 12 are larger than the base estimates in Row 2, especially for the MA samples. In particular, these results suggest that larger metropolitan areas such as Cleveland are also penalized by lower factor returns due to a lack of accessibility to one of the big-three American cities. Moreover, they also suggest that the base estimates understate the distance penalties.

The model reported in Row 13 adds to the base model the incremental population of each of the urban tiers to compare (1) the relative roles of incremental population versus incremental distances to higher tiered urban areas and (2) whether incremental population alters the market potential results described below.[24] Compared to the base case, these results suggest that changes in population of the higher-ordered urban areas do not greatly alter the distance penalty estimates nor do they affect the market potential results.[25]

Row 14 reports distance penalties when changing the base model to incorporate 1970 measures of the population variables, including using a 1969 measure of market potential, to assess whether using even deeper lags of the relevant explanatory variables affect the results. These results are comparable to those

---

[22] Using the Carbon County example in Figure 3, because Los Angeles is 1,295kms away, the incremental distance to reach Los Angeles is 1,013kms after subtracting the 282 kms to reach Denver.

[23] We conduct sensitivity analysis using the earlier 1999 definitions that mostly use 1980s commuting patterns to establish boundaries for the then existing (1990) MAs along with any new MAs that were defined during the 1990s. These results generally yielded similar findings, though the resulting distance effects and market potential effects were somewhat smaller using the older definition.

[24] For example, if the nearest/actual urban center is 25,000 people (MICRO), the next closest urban center is 670,000 population, the third closest urban center is 2.5million, then the incremental population of nearest MA is 645,000, the incremental population of a MA that is >250,000 is 0, the incremental population of a MA >500,000 is 0, and the incremental population of a MA that is at least 1.5 million is 1.83million (i.e., 2.5 million minus 670,000).

[25] We also calculated the penalties from a model that omits the lagged 1990 county population and nearest/actual total urban population from the base model in Row 2 to assess whether their inclusion affects the results (either through multicollinearity or omitted variable bias). Yet the distance-penalty results are not different from the base (not shown).

reported in Row 2, again indicating that underlying endogeneity is not affecting the distance penalty results.

In summary, estimated distance penalties for the base model are economically consequential and precisely estimated. Extensive sensitivity analysis supports the conclusions from the base model estimates, in which if anything, adding further nonlinearities indicates that the base figures are conservatively estimated. The urban hierarchy distance effects apply regardless of whether the market potential variables are included in the model, showing how more finely delineated urban access measures produce agglomeration spillovers that are distinctly different from aggregate market potential.

Market Potential Effects. Akin to Table 3, Table 4 reports the effects on median earnings and housing costs of the four market potential variables when measured at their mean levels. The market potential effects tend to be smaller than the distance penalties. For both the nonmetro and rural samples, market potential is associated with a 0 to 4% increase in median earnings. Housing costs appear to be influenced more, with the corresponding response ranging between 4 and 9% when the distance variables are included and 9-15% when they are not (ignoring the apparently misspecified log-distance results). These results are consistent with the home market profit effects of NEG, but the relatively small magnitude compared to the distance penalties indicates that the spatial placement of the urban population matters more.

In the two MA samples, the market potential effects are generally statistically insignificant, especially for small MAs. In fact, when evaluated at the mean of the market potential variables, median earnings and average housing costs often exhibit a decline in the MA samples. Indeed, the results suggest small or negative effects even when omitting the urban hierarchy distance variables, though the market potential effects appear to be overestimated for housing costs when omitting the distance terms (in Rows 8 and 11).[26]

Regarding MAs, the market potential results are generally consistent with agglomeration shadows, but not other NEG effects found in other studies. That is, after accounting for favorable effects of being closer to higher-tiered urban areas, added spatial competition reflected by greater market potential appears to reduce

---

[26] As described above, Row 14 reports the estimates when using 1970 measures of the population variables—including a 1969 measure of market potential. The market potential results are similar to those reported in Row 2 for the base model. We further consider whether underlying endogeneity in the market potential results was affecting the results by estimating an IV model that used 1950 population at specific distances as an instrument for 1989 market potential. For example, we used the 1950 population in counties located 400 to 500kms to predict 1989 market potential 400-500kms away. Illustrating that 1950 population is a strong instrument, first-stage p-values for the null hypothesis that its coefficient equals zero were better than .0001. Yet, given our expectation that the market potential results are likely positively biased, it was somewhat surprising that in the case of median earnings, the IV small MA market potential effect approximately equaled the base distance penalty, while for large MAs, the IV market potential effect was estimated to be approximately zero—rather than negative and significant. Nevertheless, the general pattern was that distance to specific urban tiers still matters more than overall measures of market potential.

factor returns in MAs with greater than 500K residents (in 1990). The relative magnitude of these impacts also suggests that while market potential affects earnings and housing costs, distance to the urban hierarchy appears to have more consequential effects. That is, it matters whether the market potential/population is concentrated into larger cities as suggested by our adaptation of the urban hierarchy approach versus a situation of having the same market potential but with population being more uniformly dispersed in multiple smaller cities (a distinction not apparent in market potential as typically calculated). Thus, agglomeration spillovers appear to have much broader impacts than just the scale of the market, and appear to be transmitted differently depending on location and position in the urban system.

4.3 Mapping the Effects of Urban Hierarchy Distance and Market Potential

The estimated total distance effects on median earnings and housing costs by county are mapped in Figures 5 and 6, with similar county estimates for market potential in Figures 7 and 8. These are derived from the base model shown in columns 5, 7, and 8 of Tables 2A and 2B. Conservatively, they use only the distance and market potential coefficients that are significant at the 10% level multiplied by the exact county distances or the exact market potential and following the discussion in footnote 16. Yet, as noted in Section 4.2, these distance penalties are likely understated due to other nonlinearities.[27]

In general, distance from higher tiers in the urban hierarchy has large negative impacts on earnings and housing costs centered in the Great Plains and the nonmetropolitan Rocky Mountain west, with more isolated strong impacts occurring in Maine, the Carolinas and central Appalachia. In the Northern Great Plains, distance penalties of 12 to 40% on median earnings and 33 to 50% on housing costs are not uncommon. By definition of central places, these distance penalties are zero in MAs with more than 1.5 million people in our base model.[28] They also are generally small in the Mid-Atlantic region and in an area roughly extending from New York City to Chicago.

Recall from row 2 of Table 4, that market potential has less influence on MA county earnings and housing costs than proximity in the urban hierarchy. Yet, for nonmetropolitan counties, it is not surprising

---

[27] Only the effects of distance to higher-tiered urban areas and to market potential are mapped. This ignores agglomeration economies associated with very local market potential effects and very local distance effects. Yet, as noted above, we expect that these localized effects are relatively small, especially in MAs (as reflected by the mixed own-market potential and own-distance effects captured by the population and square miles coefficients).

[28] Recall from Row 12 of Table 3 that the estimated distance penalties are even larger when using Chicago, Los Angeles, and New York as the highest urban tier. This would also imply that all urban areas with over 1.5 million population (except the big three) would have a distance penalty due to remoteness from these three urban areas. Take the Cleveland MA for example, when including the big three as the highest tier, Cleveland would have a median earnings distance penalty of about -10% and a corresponding housing cost penalty of -20% due to distance from Chicago (New York is slightly more distant).

that the largest *positive* market potential effects are in the eastern half of the country. For example, compared to a hypothetical county with no market potential, nonmetropolitan counties in the Mid-Atlantic region and in a band running roughly from Chicago to New York have median earnings that are about 8 to 30% greater while housing costs are about 12 to 50% greater. Conversely, with a handful of exceptions, market potential has a smaller influence in the western U.S. with its lower population densities.

Overall, the market potential effects and distance penalties are of about equal size in the East. In this case, market potential appears to have slightly larger impacts on median earnings, while distance has slightly larger impacts on housing costs. Yet in the western U.S., distance has a much greater impact on both earnings and housing costs, with discrete spatial jumps in the gradients evident. As noted above, one explanation for stronger market potential effects in the East is significantly greater population density, while remoteness is more important in the more sparsely populated West. This pattern could partly underlie the strong market potential and corresponding NEG findings in explaining spatial differences in more densely populated Western Europe, while such conclusions are not correspondingly universal in other geographic settings (especially if there are national border effects). In particular, in less densely populated areas, agglomeration economies appear to be transmitted more through urban hierarchal effects via access to higher-tiered cities. Thus, both types of impacts can be important, but their overall influence appears to depend on the spatial distribution of the surrounding population.

5. CONCLUSION

We investigated the spatial dimensions of the determinants of equilibrium differentials in land and labor costs across the urban hierarchy. Our approach drew from a number of theoretical and empirical pursuits. One, NEG emphasizes the importance of access to customers and suppliers and its impact on factor returns. Two, models in the tradition of Roback (1982) articulate wage-rent trade-offs for households and firms in their location decisions. In the latter models, the net effect of agglomeration on land rents and wages is indeterminate depending on the relative strength of labor demand influences by firms and supply responses by households. Third, geographic spillovers from agglomeration centers need to be considered inasmuch as commuting and input-out links produce influences that extend well beyond the centers of economic activity. Finally, urban hierarchical effects on the factor price gradients may reveal influences beyond those predicted by NEG market potential.

Our results show significant, and hierarchically dependent, distance penalties for both median earnings and housing costs. This pattern holds for all four community types—non-metro, rural, small MAs and large MAs. These results are robust to the addition of market potential variables as well as other specifications. Distance from urban centers inflicts a penalty on earnings and housing costs, and there is an incremental penalty for remoteness from successively higher tiers in the urban hierarchy. By examining both earnings and housing costs, we conclude that the proximity to higher tiers in the urban hierarchy is attractive for firms, which dominates any potential aversion to urban congestion or attractiveness of remoteness that households may exhibit. The market potential effects are generally less significant than the distance effects, though these effects are stronger for nonmetropolitan areas.

These cumulative distance penalties across all urban tiers show that the factor price penalties for remoteness are greater for housing costs than for median earnings and are confined to a relatively tight and consistent band of 6-9% lower earnings and 15-20% lower housing costs. Somewhat surprisingly, the cumulative penalties are somewhat larger for the urban than the rural samples, even though there would be fewer higher tiers over which the penalties are accumulated. This could suggest that firms in smaller MAs are more influenced by factor prices in larger metropolitan areas—i.e., they are more economically integrated, such that greater distance imposes a larger marginal penalty.

Maps showing visual representation of the results confirm the expected large negative impacts of remoteness in the Great Plains as well as in scattered non-metro counties generally. The price of remoteness—or the tyranny of distance—is up to 40% in median earnings, and up to 50% in terms of housing costs. Similarly the positive market potential effects are evident in the more densely populated eastern U.S., a finding consistent with the standard NEG predictions, suggesting they may be most germane in more densely populated settings including western Europe.

In summary, the overall results suggest that urban hierarchical distances are more important than market potential; it matters whether the nearby population is concentrated into large cities or more evenly distributed into smaller cities. Thus, the urban hierarchy provides a valuable framework for understanding the spatial patterns in factor price differentials. Future research is needed to better understand the nature of the economic linkages between firms across the urban hierarchy and the role of households.

References

Blomquist, Glenn, Mark Berger, and John Hoehn, 1988. "New Estimates of the Quality of Life in Urban Areas," *American Economic Review* 78, 89-107.

Brülhart, Marius and Pamina Koenig (2006). "New Economic Geography Meets Comecon," *Economics of Transition* 14(2), 245-267.

Christaller, W. 1933. "Central Places in Southern Germany." Translated from *Die Zentralen Orte in Süddeutschland* by Charles W. Baskin. Englewood Cliffs, N.J. Prentice Hall.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. (forthcoming). "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics.*

Conley, T.G., 1999. "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics* 92, 1-45.

Davis, Morris A. and Michael G. Palumbo, 2006. "The Price of Residential Land in Large U.S. Cities," Board of Governors of the Federal Reserve System (U.S.), Finance and Economics Discussion Series: 2006-25.

Davis, Donald R. and David E. Weinstein, 2003. "Market Access, Economic Geography and Comparative Advantage: An Empirical Test," *Journal of International Economics* 59(1), 1-23.

Dekle, Robert and Jonathon Eaton, 1999. "Agglomeration and Land Rents: Evidence from the Prefectures," *Journal of Urban Economics* 46(2), 200-214.

Desmet, Klaus and Esteban Rossi-Hansberg. 2007. "Spatial Growth and the Industrial Age." *NBER Working Paper #13302.*

Eaton, Jonathan and Zvi Eckstein, 1997. "Cities and Growth: Theory and Evidence from France and Japan," *Regional Science and Urban Economics* 27(4-5): 443-74.

Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica,* 70, 1741-1779.

Fingleton, Bernard. 2004. "Some Alternative Geo-economics for Europe's Regions." *Journal of Economic Geography* 36, 595-612.

Fingleton, Bernard, 2006. "The New Economic Geography versus Urban Economics: An Evaluation using Local Wage Rates in Great Britain," *Oxford Economic Papers* 58, 501-530.

Gabriel, Stuart A., Joe P. Mattey and William L.Wascher, 2003. "Compensating Differentials and Evolution in the Quality-of-Life among U.S. States," *Regional Science and Urban Economics* 33(5), 619-649.

Gabriel, Stuart A. and Stuart S. Rosenthal, 2004. "Quality of the Business Environment versus Quality of Life: Do Firms and Households Like the Same Cities?" *Review of Economics and Statistics* 86(1), 438-444.

Glaeser, Edward L., 1997. "Are Cities Dying?" *Journal of Economic Perspectives* 12(2), 139-160.

Glaeser, Edward L., 2007. "Did the Death of Distance Hurt Detroit and Help New York?" National Bureau of Economic Research Working Paper 13710.

Glaeser, Edward L. and Janet E. Kohlhase, 2004. "Cities, Regions, and the Decline of Transport Costs," *Papers in Regional Science* 83, 197-228.

Glaeser, Edward L. and David C. Maré. 2001. “Cities and Skills,” *Journal of Labor Economics,* 19, 316-342.

Gyourko, Joseph, Matthew Kahn, and Joseph Tracy, 1999. “Quality of Life and Environmental Comparisons,” in Edwin S. Mills and Paul Cheshire (eds.), *Handbook of Regional and Urban Economics*. New York: North Holland Press, pp. 1413-54.

Hanson, Gordon H., 2005. “Market Potential, Increasing Returns and Geographic Concentration,” *Journal of International Economics* 67, 1-24.

Head, Keith and Thierry Mayer, 2004a. “The Empirics of Agglomeration and Trade,” in Handbook of Regional and Urban Economics, Vol. 4, pp. 2609-69, Eds. (J.V. Henderson and J.F. Thisse). Elsevier: San Diego. Accessed at http://strategy.sauder.ubc.ca/head/Papers/neat.pdf.

_____, 2004b. “Market Potential and the Location of Japanese Investment in the European Union,” *Review of Economics and Statistics* 86(4), 959-972.

_____, 2006. “Regional Wage and Employment Responses to Market Potential in the EU,” *Regional Science and Urban Economics* 36, 573-594.

Harris, Curtis, 1954. “The Market as a Factor in the Localization of Industry in the United States*,” Annals of the Association of American Geographers* 64, 315-348.

Jacobs, J. (1969). *The Economy of Cities*. Random House: New York.

Jolliffe, Dean, 2006a. “Adjusting for Living Costs Can Change Who Is Considered Poor.” *Amber Waves* (November), Available at: http://www.ers.usda.gov/AmberWaves/November06/.

Jolliffe, Dean, 2006b. *The Cost of Living and the Geographic Distribution of Poverty.* Washington, DC: U.S. Department of Agriculture Economic Research Service Report #26. Available at: http://www.ers.usda.gov/publications/err26/err26.pdf.

Lucas, Robert E., Jr., and Esteban Rossi-Hansberg, 2002. “On the Internal Structure of Cities,” *Econometrica* 70(4), 1445-76.

Knaap, Thijs, 2006. “Trade, location, and wages in the United States,” *Regional Science and Urban Economics* 36, 595-612.

Krugman, P., 1996. “Confronting the Mystery of Urban Hierarchy,” *Journal of the Japanese and International Economies* 10(4), 399-418.

Mulligan, Gordon F., 1984. “Agglomeration and Central Place Theory,” *Interregional Science Review* 9(1), 1-42.

Ottaviano, Gianmarco and Dino Pinelli, 2006. “Market Potential and Productivity: Evidence from Finnish Regions,” *Regional Science and Urban Economics* 36, 636-657.

Partridge, Mark D. and Dan S. Rickman, 2006. *The Geography of American Poverty: Is There a Role for Place Based Policy?* Kalamazoo, MI: W.E. Upjohn Employment Institute.

Peiser**,** Richard B. and Lawrence B. Smith, 1985. “Homeownership Returns, Tenure Choice and Inflation,” *American Real Estate and Urban Economics Association Journal* 13(4), 343-360.

Rappaport, Jordan, 2004. “*Moving to nice weather,*” RWP 03-07. Research Division. Federal Reserve

Bank of Kansas City.

Redding, S. and A. Venables, 2004. "Economic Geography and International Inequality," *Journal of International Economics* 62(1), 53-82.

Roback, J., 1982. "Wages, Rents, and the Quality of Life," *Journal of Political Economy,* 90, 1257-1278.

Rosenthal, Stuart S. and William C. Strange, 2001. "The Determinants of Agglomeration," *Journal of Urban Economics* 50, 191-229.

_____, 2003. "Geography, Industrial Organization, and Agglomeration," *Review of Economics and Statistics* 85(2), 377-93.

_____, 2006. "The Attenuation of Human Capital Spillovers," unpublished manuscript.

Sen, Ashish and Tony E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior.* New York: Springer.

U.S. Office of Management and Budget. 2000. "Standards for Defining Metropolitan and Micropolitan Statistical Areas; Notice." *Federal Register* (December 27), 65: 82,228-82,238.

Table 1: Correlation between Urban Hierarchy Distances and Market Potential Measures*

	Aggregate Income 100-200 Kilometers	Aggregate Income 200-300 Kilometers	Aggregate Income 300-400 Kilometers	Aggregate Income 400-500 Kilometers
Distance to Nearest Urban Area	-0.25	-0.25	-0.22	-0.20
Distance to Nearest Metropolitan Area	-0.25	-0.25	-0.22	-0.19
Incremental Distance to MA>250K	-0.25	-0.26	-0.28	-0.28
Incremental Distance to MA>250K and <500K	-0.18	-0.10	-0.10	-0.07
Incremental Distance to MA > 1.5 million	-0.19	-0.16	-0.12	-0.10

*Market potential is measured by aggregate household income within the reported concentric rings surrounding the county following Hanson (2005).

Table 2A: Dependent variable: log(median earnings 1999 in $)

Variables	Distance				Base = Distance + Market Potential			
	Non-metro	Rural	Small MA≤250k	Large MA>250k	Non-metro	Rural	Small MA≤250k	Large MA>250k
Intercept	10.115*** (72.36)	10.144*** (71.20)	9.943*** (32.49)	10.070*** (30.81)	10.101*** (74.29)	10.135*** (71.90)	9.978*** (33.91)	10.218*** (33.94)
Dist to nearest urban center	-7.7E-04*** (-6.19)	-9.3E-04*** (-6.23)	N	N	-6.8E-04*** (-5.46)	-8.5E-04*** (-5.60)	N	N
Dist to the center of own metro	N	N	3.4E-04 (0.95)	-2.0E-03*** (-4.65)	N	N	4.3E-04 (1.22)	-2.1E-03*** (-5.59)
Inc dist to a MA	-3.0E-04*** (-2.87)	-1.8E-04 (-1.60)	N	N	-2.1E-04* (-1.94)	-1.3E-04 (-1.18)	N	N
Inc dist to MA>250k	-2.5E-04*** (-4.10)	-2.0E-04*** (-3.20)	-5.1E-04*** (-4.01)	N	-1.8E-04*** (-2.85)	-1.6E-04*** (-2.60)	-4.8E-04*** (-3.77)	N
Inc dist to MA>500k	-1.4E-04** (-2.07)	-1.4E-04* (-1.83)	-4.0E-04*** (-3.61)	-1.1E-03*** (-7.61)	-6.1E-05 (-0.87)	-9.6E-05 (-1.21)	-3.8E-04*** (-3.46)	-1.1E-03*** (-7.62)
Inc dist to MA>1500k	-8.4E-05* (-1.67)	-1.2E-04** (-2.50)	-3.9E-04*** (-4.24)	-5.0E-04*** (-3.19)	-2.1E-05 (-0.43)	-7.4E-05* (-1.67)	-3.8E-04*** (-3.91)	-5.2E-04*** (-3.36)
Market potential within 100-200 km 1989	N	N	N	N	4.6E-07*** (4.59)	2.8E-07** (2.17)	6.9E-08 (0.41)	-4.15E-08 (-0.25)
Market potential within 200-300 km 1989	N	N	N	N	7.2E-09 (0.12)	2.4E-08 (0.29)	1.3E-07 (0.97)	-2.7E-07 (-1.34)
Market potential within 300-400 km 1989	N	N	N	N	1.6E-07** (2.20)	1.0E-07 (1.10)	-4.6E-08 (-0.24)	-5.3E-09 (-0.02)
Market potential within 400-500 km 1989	N	N	N	N	1.1E-07 (1.26)	1.4E-07 (1.37)	-4.9E-08 (-0.35)	-2.2E-07 (-1.52)
County Population 1990	8.6E-07*** (3.99)	2.6E-06*** (7.33)	1.0E-07 (1.21)	-2.7E-08** (-2.04)	7.2E-07*** (3.69)	2.6E-06*** (6.98)	1.2E-07 (1.39)	-2.8E-08*** (-2.37)
Pop of nearest/actual urban center 1990	7.4E-08* (1.85)	7.0E-08** (2.06)	8.3E-08 (0.47)	1.0E-08 (1.04)	7.2E-08* (1.76)	7.0E-08** (1.96)	7.0E-08 (0.38)	8.0E-09 (0.79)
County area (sq miles)	-8.3E-06** (-2.36)	-1.7E-06 (-0.40)	-1.3E-05* (-1.72)	-9.5E-06 (-1.57)	-7.3E-06** (-2.16)	-9.0E-07 (-0.21)	-1.4E-05* (-1.80)	-9.1E-06 (-1.61)
Amenities/Ocean	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.36	0.41	0.31	0.45	0.37	0.41	0.30	0.46
Sample size	1972	1300	416	640	1972	1300	416	640
F-stats								
All distance vars = 0	16.32***	13.80***	9.20***	37.36***	10.68***	9.65***	8.04***	38.96***
Incremental dist = 0	10.76***	5.71***	9.47***	54.38***	4.67***	2.86***	7.52***	56.61***
Market potentials = 0	N	N	N	N	7.78***	2.42***	0.28	1.80

Notes: Robust t-statistics following Conley (1999) and Rappaport (2004) are in the parentheses. A ***, **, and * indicate significant at 1%, 5%, and 10% levels respectively. Y=included, N=not included. Rural omits 672 micropolitan area counties from the non metro sample. For variable descriptions, see the text and Appendix Table 1. All models further include the following Amenities/Ocean variables: Jan sun hours, Jan temp, July humidity, July temp; USDA topography score, % water area in the county, and three indicators for being located within 50kms of one of the Great Lakes, Pacific Ocean, and Atlantic Ocean.

Table 2B: Dependent variable: log(weighted average median house rent 2000 in $/month)

Variables	**Distance**				**Base = Distance + Market Potential**			
	Non-metro	Rural	Small MA≤250k	Large MA>250k	Non-metro	Rural	Small MA≤250k	Large MA>250k
Intercept	8.373***	8.486***	7.161***	7.186***	8.359***	8.462***	7.246***	7.513***
	(25.66)	(21.60)	(15.24)	(12.89)	(26.35)	(21.87)	(16.84)	(15.70)
Dist to nearest urban center	-2.3E-03***	-1.7E-03***	N	N	-2.2E-03***	-1.4E-03***	N	N
	(-9.26)	(-4.96)			(-8.50)	(-4.14)		
Dist to the center of own metro	N	N	-2.8E-03***	-4.0E-03***	N	N	-2.7E-03***	-4.1E-03***
			(-3.58)	(-5.55)			(-3.42)	(-6.54)
Inc dist to a MA	-6.5E-04***	-5.2E-04*	N	N	-4.5E-04**	-3.3E-04	N	N
	(-3.10)	(-1.92)			(-2.16)	(-1.27)		
Inc dist to MA>250k	-6.9E-04***	-6.3E-04***	-1.1E-03***	N	-5.5E-04***	-5.0E-04***	-1.1E-03***	N
	(-5.27)	(-4.19)	(-5.42)		(-4.15)	(-3.41)	(-5.16)	
Inc dist to MA>500k	-5.3E-04***	-6.0E-04***	-5.4E-04***	-1.7E-03***	-4.0E-04***	-4.5E-04***	-5.3E-04***	-1.7E-03***
	(-3.75)	(-3.45)	(-2.63)	(-6.89)	(-3.00)	(-2.81)	(-2.46)	(-6.80)
Inc dist to MA>1500k	-2.3E-04**	-2.5E-04**	-4.8E-04***	-6.8E-04***	-1.2E-04	-1.0E-04	-5.2E-04***	-7.1E-04***
	(-2.25)	(-2.00)	(-3.20)	(-3.00)	(-1.35)	(-0.92)	(-3.27)	(-3.22)
Market potential within 100-200 km 1989	N	N	N	N	6.8E-07***	9.3E-07***	-1.9E-08	-1.2E-07
					(3.68)	(3.53)	(-0.08)	(-0.42)
Market potential within 200-300 km 1989	N	N	N	N	2.1E-07	1.8E-07	5.5E-08	-4.9E-07
					(1.42)	(0.88)	(0.18)	(-1.43)
Market potential within 300-400 km 1989	N	N	N	N	3.7E-07***	6.2E-07***	5.8E-08	-2.7E-07
					(2.55)	(3.44)	(0.19)	(-0.80)
Market potential within 400-500 km 1989	N	N	N	N	-1.6E-08	1.8E-07	-2.8E-07	-4.0E-07
					(-0.11)	(1.05)	(-1.07)	(-1.44)
County Population 1990	3.0E-06***	7.5E-06***	7.5E-07***	-6.8E-08***	2.8E-06***	7.3E-06***	7.6E-07***	-6.9E-08***
	(6.79)	(6.34)	(3.26)	(-2.98)	(6.03)	(6.16)	(3.20)	(-3.31)
Pop of nearest/actual urban center 1990	1.9E-07**	1.6E-07*	-1.1E-07	2.4E-08*	2.0E-07**	1.7E-07*	-1.6E-07	2.1E-08
	(2.16)	(1.83)	(-0.52)	(1.75)	(2.23)	(1.93)	(-0.84)	(1.38)
County area (sq miles)	-4.9E-06	-7.6E-06	-2.9E-06	-2.2E-05***	-3.4E-06	-4.1E-06	-4.7E-06	-2.1E-05***
	(-0.76)	(-0.71)	(-0.28)	(-2.82)	(-0.53)	(-0.40)	(-0.44)	(-3.07)
Amenities/Ocean	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.64	0.63	0.61	0.59	0.64	0.64	0.61	0.59
Sample size	1972	1300	416	640	1972	1300	416	640
F-stats								
All distance vars = 0	47.55***	17.54***	11.64***	36.53***	36.67***	10.35***	10.26***	38.26***
Incremental dist = 0	24.26***	13.57***	14.52***	47.77***	12.67***	7.21***	12.09***	50.14***
Market potentials = 0	N	N	N	N	8.36***	7.51***	0.56	2.28**

Notes: Robust t-statistics following Conley (1999) and Rappaport (2004) are in the parentheses. A ***, **, and * indicate significant at 1%, 5%, and 10% levels respectively. Y=included, N=not included. Rural omits 672 micropolitan area counties from the non metro sample. For variable descriptions, see the text and Appendix Table 1. All models further include the following Amenities/Ocean variables: Jan sun hours, Jan temp, July humidity, July temp; USDA topography score, % water area in the county, and three indicators for being located within 50kms of one of the Great Lakes, Pacific Ocean, and Atlantic Ocean.

Table 3: Distance Penalties Evaluated at the Mean Distances

Models	Non-metro		Rural		Small MA ≤ 250,000 pop		Large MA > 250,000 pop	
	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)
Median Earnings								
(1) Dist	-7.58 (-8.03)	-7.58 (-8.03)	-8.35 (-8.02)	-9.05 (-7.91)	-8.85 (-5.28)	-8.85 (-5.28)	-8.53 (-8.89)	-8.53 (-8.89)
(2) Base = (1) + mkt potential	-5.04 (-6.33)	-5.47 (-5.20)	-6.73 (-6.44)	-7.67 (-6.02)	-8.51 (-4.70)	-8.51 (-4.70)	-8.82 (-9.08)	-8.82 (-9.08)
(3) Base + sq distance	-9.52 (-7.00)	-9.70 (-6.28)	-16.85 (-8.68)	-17.42 (-8.33)	-15.20 (-5.95)	-15.01 (-6.20)	-12.93 (-7.32)	-12.33 (-8.34)
(4) Base+sq dist+cubic dist	-12.58 (-7.79)	-10.81 (-5.67)	-25.84 (-4.70)	-22.13 (-7.04)	-19.02 (-4.82)	-17.66 (-5.47)	-15.02 (-5.47)	-13.91 (-7.61)
(5) Base with log(distance)	-3.08 (-4.49)	-2.92 (-3.81)	-13.67 (-8.08)	-14.00 (-8.17)	-6.61 (-4.72)	-6.61 (-4.72)	-6.05 (-7.92)	-6.05 (-7.92)
(6) Full90 = (2) + 1990 vars	-7.02 (-10.11)	-7.02 (-10.11)	-8.42 (-8.68)	-8.42 (-8.68)	-6.22 (-5.86)	-6.22 (-5.86)	-4.06 (-7.87)	-4.06 (-7.87)
(7) Full90 – market potential	-7.73 (-12.11)	-7.73 (-12.11)	-8.51 (-9.56)	-8.51 (-9.56)	-6.31 (-6.41)	-6.31 (-6.41)	-3.86 (-7.48)	-3.86 (-7.48)
(8) Full90 – distance	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(9) Full00 = (2) + 2000 vars	-6.11 (-9.41)	-6.11 (-9.41)	-7.03 (-7.57)	-7.03 (-7.57)	-4.77 (-5.53)	-4.77 (-5.53)	-3.84 (-8.83)	-3.84 (-8.83)
(10) Full00 – market potential	-7.17 (-12.09)	-7.17 (-12.09)	-7.72 (-9.10)	-7.72 (-9.10)	-5.09 (-6.29)	-5.09 (-6.29)	-3.69 (-8.49)	-3.69 (-8.49)
(11) Full00 – distance	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(12) Base with inc dist to BIG3	-10.26 (-5.38)	-11.14 (-5.02)	-12.38 (-5.23)	-13.03 (-5.27)	-19.51 (-3.72)	-19.51 (-3.72)	-20.32 (-5.05)	-20.32 (-5.05)
(13) Base + inc MA pop	-4.58 (-5.60)	-4.41 (-3.88)	-5.71 (-5.44)	-6.59 (-4.82)	-8.63 (-4.37)	-8.63 (-4.37)	-8.86 (-9.07)	-8.86 (-9.07)
(14) Base with 1970 vars	-5.44 (-6.89)	-5.94 (-5.76)	-7.31 (-7.01)	-8.46 (-6.70)	-8.47 (-4.83)	-8.47 (-4.83)	-9.24 (-9.73)	-9.24 (-9.73)
Housing Cost								
(1) Dist	-19.81 (-12.19)	-19.81 (-12.19)	-19.59 (-8.91)	-19.59 (-8.91)	-14.49 (-6.41)	-14.49 (-6.41)	-12.07 (-7.97)	-12.07 (-7.97)
(2) Base = (1) + mkt potential	-15.53 (-10.69)	-16.46 (-9.57)	-13.18 (-6.91)	-15.17 (-6.13)	-14.65 (-6.03)	-14.65 (-6.03)	-12.45 (-8.15)	-12.45 (-8.15)
(3) Base + sq distance	-27.58 (-12.37)	-27.32 (-11.45)	-28.38 (-8.46)	-30.29 (-7.77)	-24.73 (-7.42)	-24.29 (-7.67)	-17.37 (-6.20)	-16.80 (-7.17)
(4) Base+sq dist+cubic dist	-31.06 (-9.05)	-30.96 (-10.64)	-31.22 (-6.04)	-37.11 (-6.43)	-32.66 (-8.36)	-32.99 (-8.30)	-21.85 (-5.04)	-19.48 (-6.70)
(5) Base with log(distance)	-9.51 (-8.15)	-9.29 (-7.13)	-18.95 (-6.45)	-21.07 (-6.26)	-12.71 (-6.80)	-12.71 (-6.80)	-8.24 (-6.70)	-8.24 (-6.70)
(6) Full90 = (2) + 1990 vars	-14.03 (-11.93)	-14.93 (-10.73)	-16.13 (-9.07)	-16.76 (-8.22)	-12.00 (-6.62)	-12.00 (-6.62)	-7.14 (-6.42)	-7.14 (-6.42)
(7) Full90 – market potential	-18.01 (-14.47)	-18.01 (-14.47)	-19.54 (-10.66)	-19.54 (-10.66)	-11.47 (-6.81)	-11.47 (-6.81)	-6.65 (-6.04)	-6.65 (-6.04)
(8) Full90 – distance	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(9) Full00 = (2) + 2000 vars	-8.27 (-7.99)	-9.33 (-6.92)	-10.83 (-6.23)	-11.46 (-5.78)	-6.36 (-3.69)	-6.36 (-3.69)	-5.16 (-4.85)	-5.16 (-4.85)
(10) Full00 – market potential	-11.59 (-9.44)	-11.59 (-9.44)	-13.27 (-7.33)	-13.27 (-7.33)	-5.87 (-3.64)	-5.87 (-3.64)	-4.79 (-4.53)	-4.79 (-4.53)
(11) Full00 – distance	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(12) Base with inc dist to BIG3	-18.13 (-10.86)	-22.71 (-6.28)	-23.03 (-7.07)	-24.37 (-5.21)	-29.28 (-4.28)	-29.28 (-4.28)	-34.39 (-5.86)	-34.39 (-5.86)
(13) Base + inc MA pop	-14.67 (-9.51)	-15.33 (-8.22)	-12.50 (-6.19)	-14.45 (-5.44)	-11.86 (-5.41)	-12.95 (-4.97)	-12.49 (-8.17)	-12.49 (-8.17)
(14) Base with 1970 vars	-17.67 (-12.29)	-18.64 (-11.00)	-17.00 (-7.87)	-17.98 (-7.31)	-14.90 (-6.29)	-14.90 (-6.29)	-13.10 (-8.79)	-13.10 (-8.79)

t-stat refers to the significance of the untransformed linear combination obtained from the ”lincom” command in STATA.

<u>Median Earnings models description</u>:

(1) Dist: includes 5 distance variables, plus 1990 county pop, 1990 pop in the nearest urban center, January sun hours, January temp, July humidity, July temp, topography score, % water area, proximity to Great Lakes, Pacific Ocean, and Atlantic Ocean, county area and state fixed effects.

(2) Base: (1) + 4 market potential variables defined as 1989 aggregate household income within 100-200 km, 200-300 km, 300-400 km, and 400-500 km rings from the county centroid; (3) = (2) + 5 quadratic distance terms; (4) = (2) + 5 quadratic and 5 cubic distance terms; (5) = similar to (2) but distance variables are in logs;

(6) Full90: (2) plus 5 ethnicity vars, 6 age-distribution vars, 4 education vars, % female, % married, and % with a work disability, all measured in 1990

(7) = (6) – market potential variables defined in (2); (8) = (6) – 5 distance variables; (9) Full00: (2) plus same set of additional variables in (6) measured in 2000

(10) = (9) – market potential variables defined in (2); (11) = (9) – 5 distance variables; (12) = (2) with incremental distance to the nearest BIG3 cities (New York, Chicago, Los Angeles); (13) = (2) + 4 incremental MA pop 1990: inc pop of the nearest MA, inc pop of MA>250,000, inc pop of MA>500,000, and inc pop of MA>1.5 mill. See footnote 24.

(14) = similar to (2) but pop variables are measured in 1970 and 1989 agg hh inc variables are replaced with 1969 BEA personal incomes within corresponding rings.

<u>Housing Cost models</u>: Same set of variables as above except Full90 and Full00, which are explained as follows: Full90: (2) plus age of housing units, shares of 1-5 bedrooms out of total rooms, share of mobile units, share of complete plumbing and share of complete kitchen facilities, all measured in 1990. Full00: (2) plus median number of rooms, age of housing units, shares of 1-5 bedrooms out of total rooms, share of mobile units, share of complete plumbing and share of complete kitchen facilities, all measured in 2000.

Table 4: Market Potential Effects at the Mean Market Potential

Models	Non-metro		Rural		Small MA ≤ 250,000 pop		Large MA > 250,000 pop	
	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)	Coeff sig (10%) Mean (t-stat)	All coeff Mean (t-stat)
Median Earnings								
(1) Dist	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(2) Base = (1) + mkt potential	3.24 (4.81)	4.44 (3.68)	0.94 (2.52)	3.30 (2.52)	0.00 (0)	0.38 (0.1)	0.00 (0)	-5.83 (-1.76)
(3) Base + sq distance	3.21 (4.75)	4.43 (3.68)	0.76 (2.06)	2.73 (2.11)	0.00 (0)	2.29 (0.61)	0.00 (0)	-4.12 (-1.27)
(4) Base+sq dist+cubic dist	2.92 (4.34)	4.03 (3.36)	0.76 (2.06)	2.56 (1.98)	1.77 (1.25)	3.15 (0.83)	0.00 (0)	-4.31 (-1.35)
(5) Base with log(distance)	5.61 (6.27)	6.03 (5.17)	0.98 (2.77)	3.44 (2.75)	4.29 (2.16)	4.45 (1.20)	0.00 (0)	-3.11 (-0.90)
(6) Full90 = (2) + 1990 vars	0.75 (3.28)	1.41 (1.78)	0.00 (0)	-0.16 (-0.16)	0.00 (0)	-0.61 (-0.28)	-4.79 (-4.56)	-5.70 (-3.46)
(7) Full90 – market potential	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(8) Full90 – distance	3.19 (4.89)	3.69 (4.71)	0.97 (3.59)	2.29 (2.38)	1.27 (1.91)	1.24 (0.55)	-3.68 (-3.31)	-3.94 (-2.25)
(9) Full00 = (2) + 2000 vars	1.06 (4.97)	1.67 (2.26)	0.67 (2.51)	0.76 (0.81)	-1.76 (-2.23)	-0.93 (-0.54)	-3.96 (-4.33)	-4.22 (-2.97)
(10) Full00 – market potential	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(11) Full00 – distance	3.26 (5.77)	3.70 (5.08)	2.34 (4.15)	2.78 (3.03)	-0.09 (-0.09)	0.62 (0.35)	-0.63 (-0.85)	-2.40 (-1.57)
(12) Base with inc dist to BIG3	1.42 (2.98)	2.30 (1.66)	0.00 (0)	1.56 (1.07)	0.00 (0)	-5.37 (-1.22)	-7.37 (-3.37)	-10.31 (-2.91)
(13) Base + inc MA pop	2.89 (4.14)	3.89 (3.18)	0.89 (2.07)	2.99 (2.24)	0.00 (0)	-0.50 (-0.13)	0.00 (0)	-5.45 (-1.62)
(14) Base with 1970 vars	3.22 (5.36)	4.57 (4.30)	0.82 (2.42)	3.19 (2.83)	2.17 (1.69)	4.11 (1.20)	0.00 (0)	-4.30 (-1.45)
Housing Cost								
(1) Dist	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(2) Base = (1) + mkt potential	6.03 (5.07)	7.30 (3.44)	8.48 (5.37)	11.55 (4.19)	0.00 (0)	-2.29 (-0.45)	0.00 (0)	-13.28 (-2.60)
(3) Base + sq distance	6.44 (5.49)	8.04 (3.84)	7.78 (5.01)	10.66 (3.94)	0.00 (0)	0.74 (0.15)	0.00 (0)	-11.41 (-2.25)
(4) Base+sq dist+cubic dist	6.13 (5.20)	7.38 (3.51)	7.64 (4.89)	10.30 (3.79)	0.00 (0)	4.12 (0.79)	0.00 (0)	-11.61 (-2.30)
(5) Base with log(distance)	11.58 (7.71)	12.65 (6.00)	9.41 (6.10)	13.36 (4.99)	2.15 (1.41)	4.01 (0.78)	0.00 (0)	-9.50 (-1.75)
(6) Full90 = (2) + 1990 vars	7.43 (6.00)	8.79 (5.15)	6.77 (5.23)	9.18 (4.07)	0.00 (0)	-6.13 (-1.69)	-7.79 (-3.52)	-12.42 (-3.59)
(7) Full90 – market potential	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(8) Full90 – distance	14.78 (8.71)	14.78 (8.71)	14.53 (6.61)	14.53 (6.61)	0.00 (0)	-0.51 (-0.13)	-4.05 (-2.62)	-9.03 (-2.49)
(9) Full00 = (2) + 2000 vars	4.47 (5.25)	6.62 (4.34)	7.18 (4.21)	6.83 (3.35)	-2.60 (-1.71)	-6.79 (-2.09)	-8.50 (-4.10)	-12.50 (-3.84)
(10) Full00 – market potential	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(11) Full00 – distance	9.64 (6.53)	9.64 (6.53)	9.63 (5.87)	10.07 (5.18)	0.00 (0)	-4.08 (-1.27)	-4.35 (-3.04)	-10.57 (-3.18)
(12) Base with inc dist to BIG3	4.73 (3.58)	4.55 (1.87)	6.83 (4.04)	7.97 (2.63)	0.00 (0)	-10.39 (-1.75)	-17.13 (-4.10)	-21.89 (-4.13)
(13) Base + inc MA pop	5.59 (4.54)	6.65 (3.08)	8.55 (5.20)	11.61 (4.13)	0.00 (0)	-4.87 (-0.98)	0.00 (0)	-11.97 (-2.31)
(14) Base with 1970 vars	5.73 (5.34)	7.47 (3.93)	7.21 (5.05)	10.51 (4.38)	0.00 (0)	1.81 (0.38)	0.00 (0)	-11.15 (-2.44)

t-stat refers to the significance of the untransformed linear combination obtained from the ”lincom” command in STATA.

Median Earnings models description:

(1) Dist: includes 5 distance variables, plus 1990 county pop, 1990 pop in the nearest urban center, January sun hours, January temp, July humidity, July temp, topography score, % water area, proximity to Great Lakes, Pacific Ocean, and Atlantic Ocean, county area and state fixed effects.

(2) Base: (1) + 4 market potential variables defined as 1989 aggregate household income within 100-200 km, 200-300 km, 300-400 km, and 400-500 km rings from the county centroid; (3) = (2) + 5 quadratic distance terms; (4) = (2) + 5 quadratic and 5 cubic distance terms; (5) = similar to (2) but distance variables are in logs;

(6) Full90: (2) plus 5 ethnicity vars, 6 age-distribution vars, 4 education vars, % female, % married, and % with a work disability, all measured in 1990

(7) = (6) – market potential variables defined in (2); (8) = (6) – 5 distance variables; (9) Full00: (2) plus same set of additional variables in (6) measured in 2000

(10) = (9) – market potential variables defined in (2); (11) = (9) – 5 distance variables; (12) = (2) with incremental distance to the nearest BIG3 cities (New York, Chicago, Los Angeles); (13) = (2) + 4 incremental MA pop 1990: inc pop of the nearest MA, inc pop of MA>250,000, inc pop of MA>500,000, and inc pop of MA>1.5 mill. See footnote 24.

(14) = similar to (2) but pop variables are measured in 1970 and 1989 agg hh inc variables are replaced with 1969 BEA personal incomes within corresponding rings.

Housing Cost models: Same set of variables as above except Full90 and Full00, which are explained as follows: Full90: (2) plus age of housing units, shares of 1-5 bedrooms out of total rooms, share of mobile units, share of complete plumbing and share of complete kitchen facilities, all measured in 1990. Full00: (2) plus median number of rooms, age of housing units, shares of 1-5 bedrooms out of total rooms, share of mobile units, share of complete plumbing and share of complete kitchen facilities, all measured in 2000.

Figure 1: Distance Effect on Spatial Equilibrium Distribution of Wages and Rent

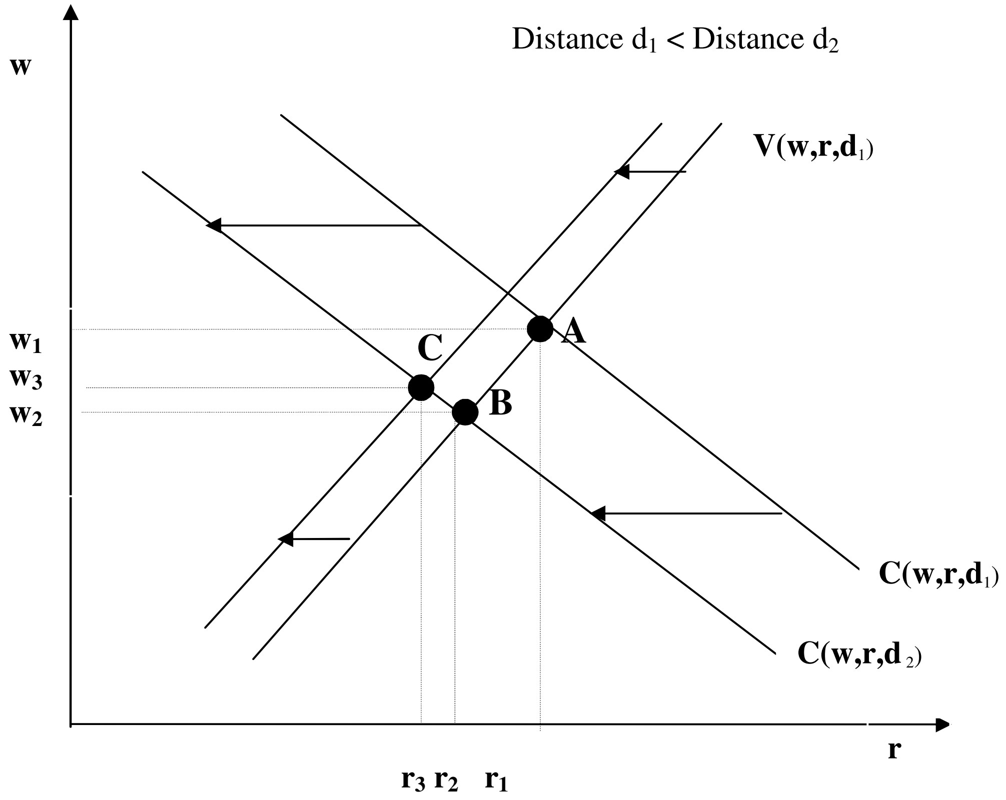
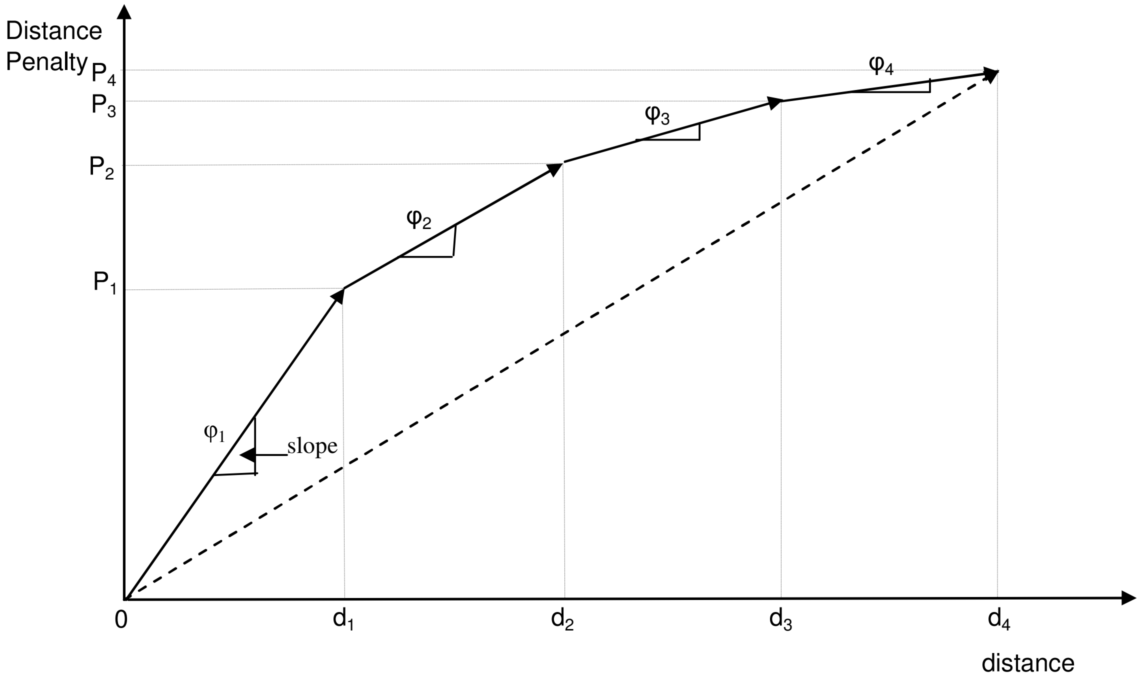
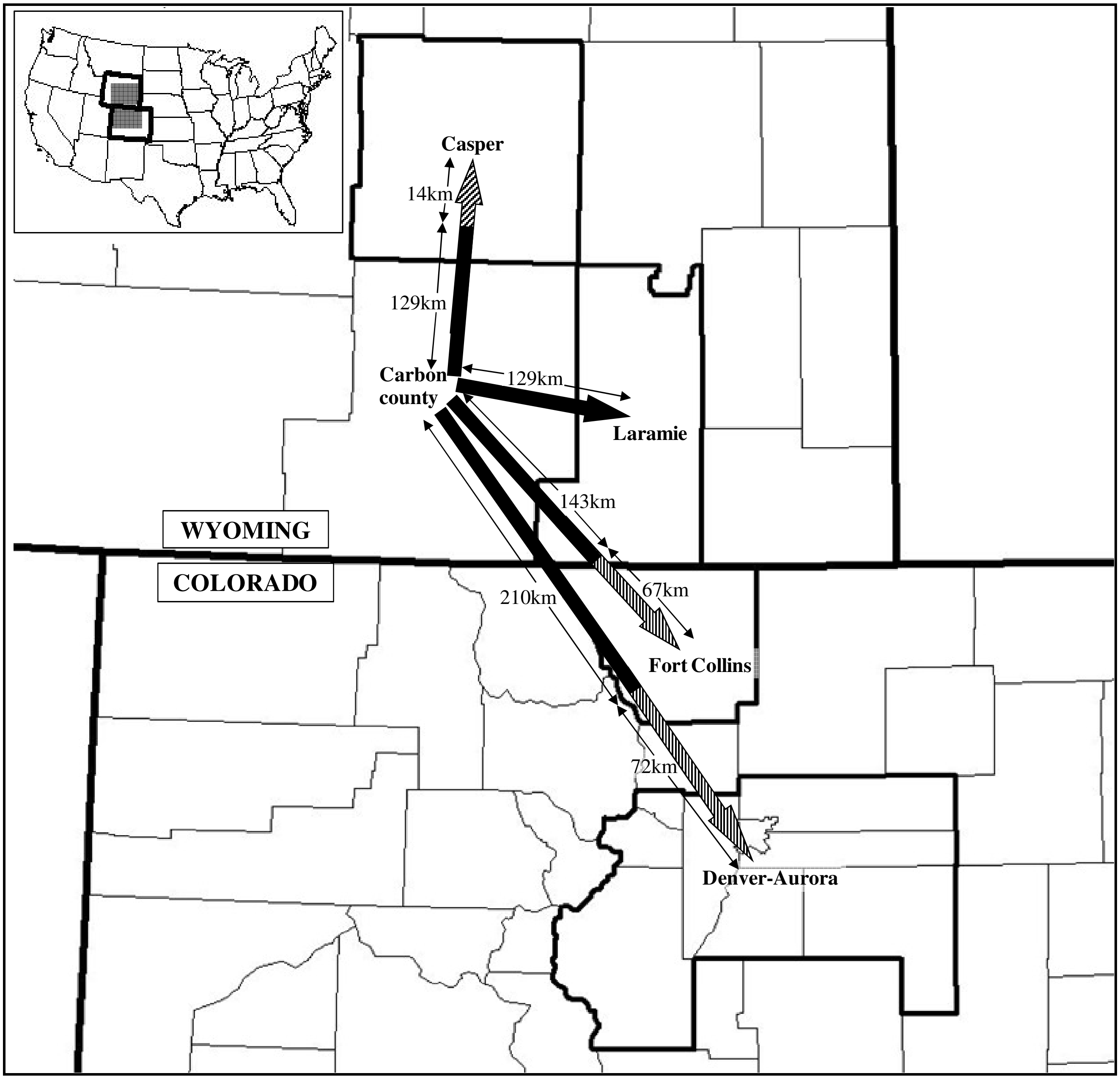


Figure 2: Illustration of Urban Hierarchy Distance Penalties



Note. The φ terms represent the slope or marginal cost of each specific distance segment.

Figure 3: Empirical Example of Urban Hierarchy Proximity: Carbon County, Wyoming



Notes: Carbon County Wyoming is a rural county. Its nearest urban area is Laramie, a micropolitan area 129kms away. The nearest metropolitan area is Casper, the nearest metropolitan area of at least 250,000 population in 1990 is Ft. Collins, and the nearest metropolitan area of at least 500,000 population and/or 1.5 million population (in 1990) is Denver. The black portion of the arrows shows the distance to the immediately lower tier and the grey cross-hatched arrow shows the incremental distance we control for. For example, Casper is 143kms, which is decomposed into the 129kms that was the distance to Laramie (the lower tiered city) and 14kms that it is *incrementally* farther than Laramie. See the text for more details.

Figure 4a: Distance and Nonmetropolitan County Median Annual Earnings ($) 1999*

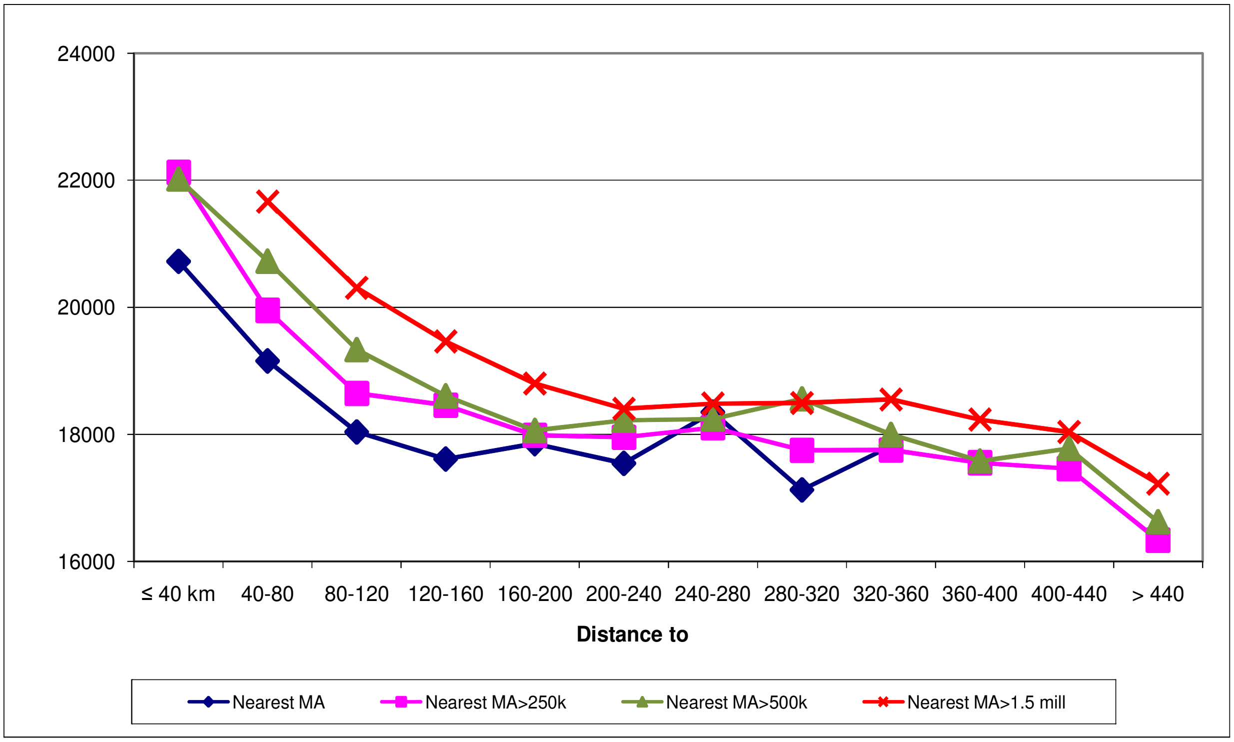
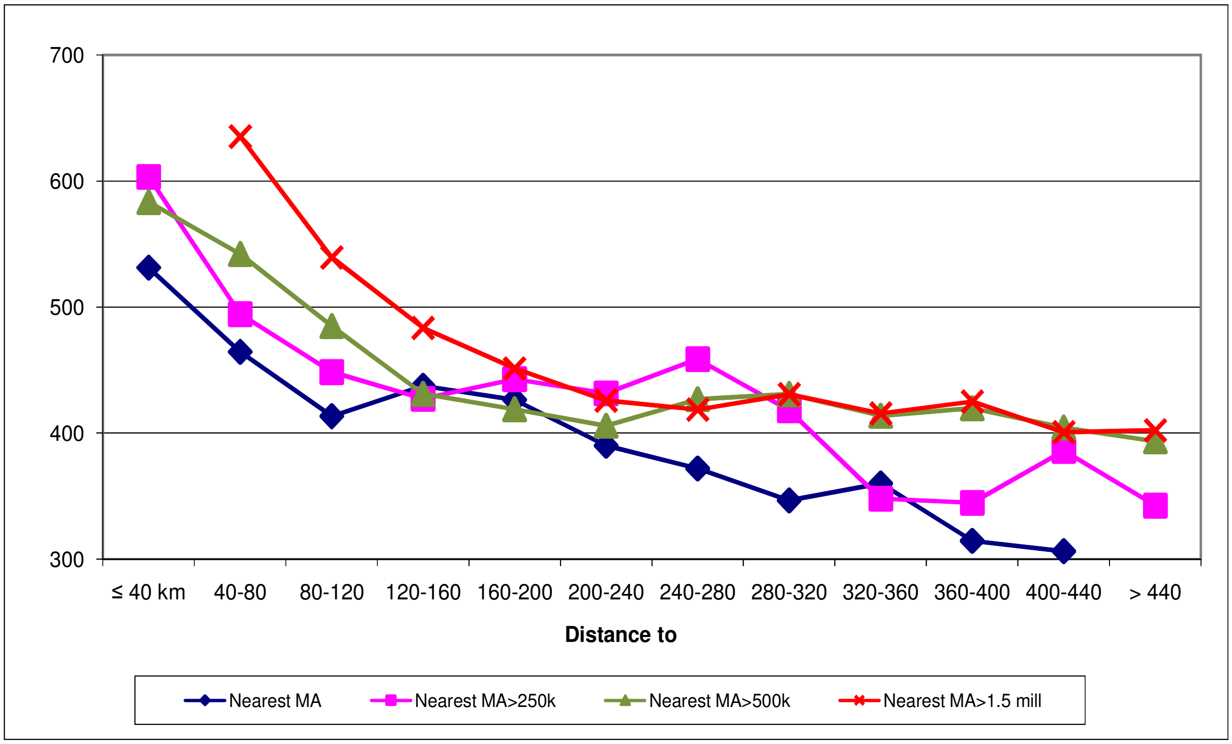
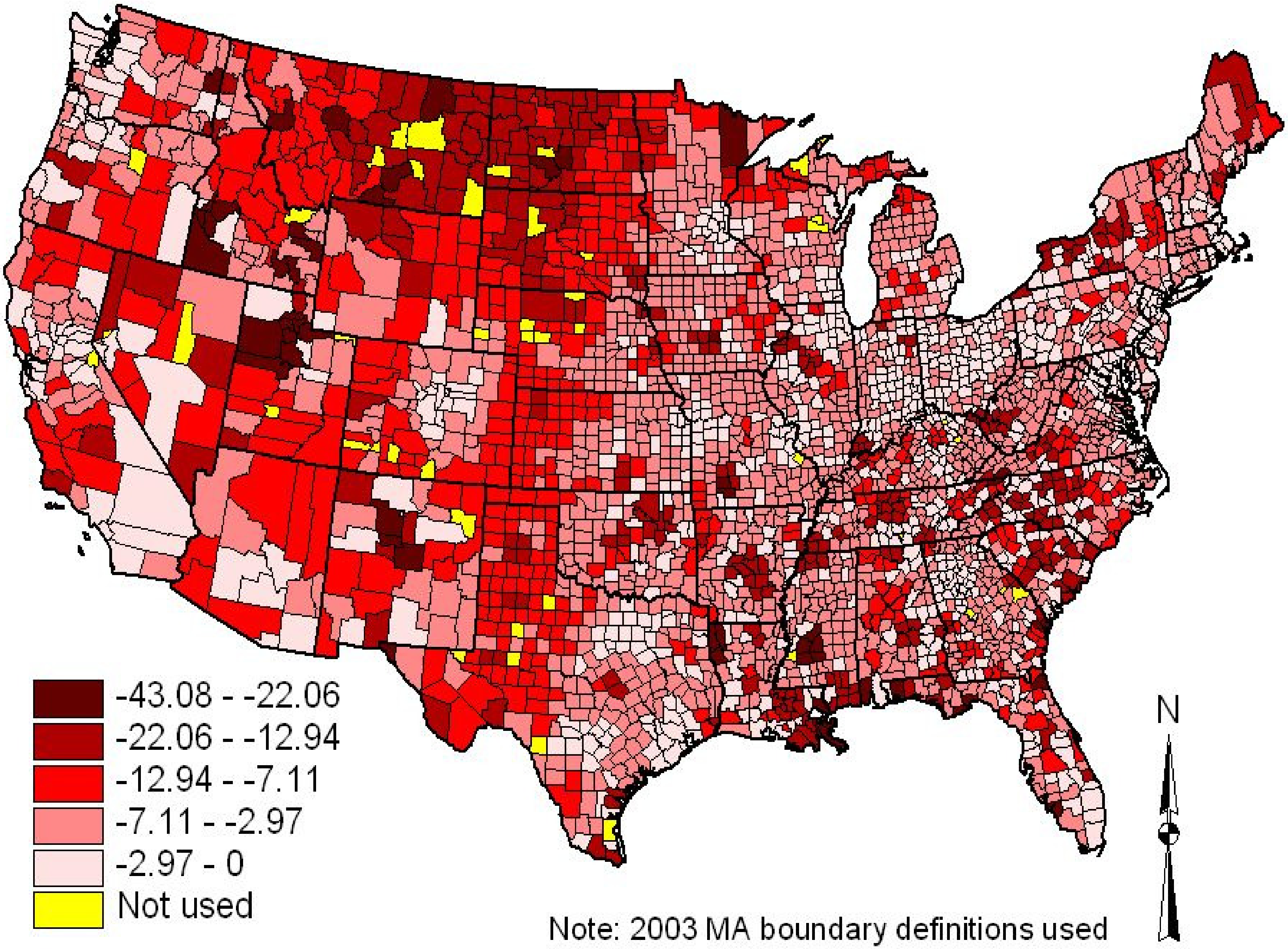


Figure 4b: Distance and Nonmetropolitan County Median Housing Costs ($/month) 2000



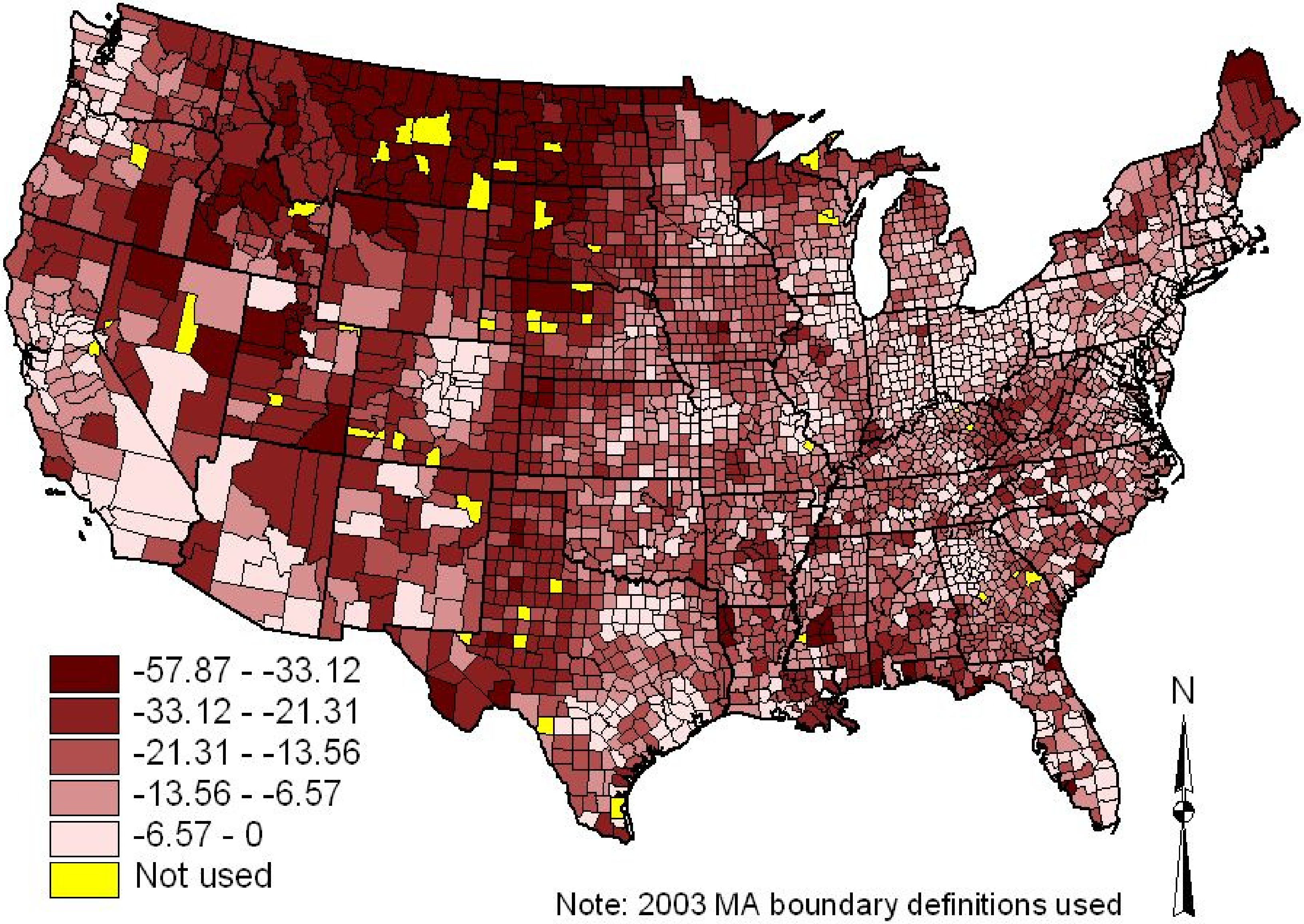
*The figures report the relationship between median earnings and median housing costs in nonmetropolitan areas with respect to distance to various higher-order urban areas. These are defined as the nearest MA of any size, nearest MA>250,000 population, nearest MA>500,000 population, and the nearest MA>1.5 million population.

Figure 5: Distance Penalties (%) for Median Earnings 1999



Note for reviewers: Color maps are in the pdf version.

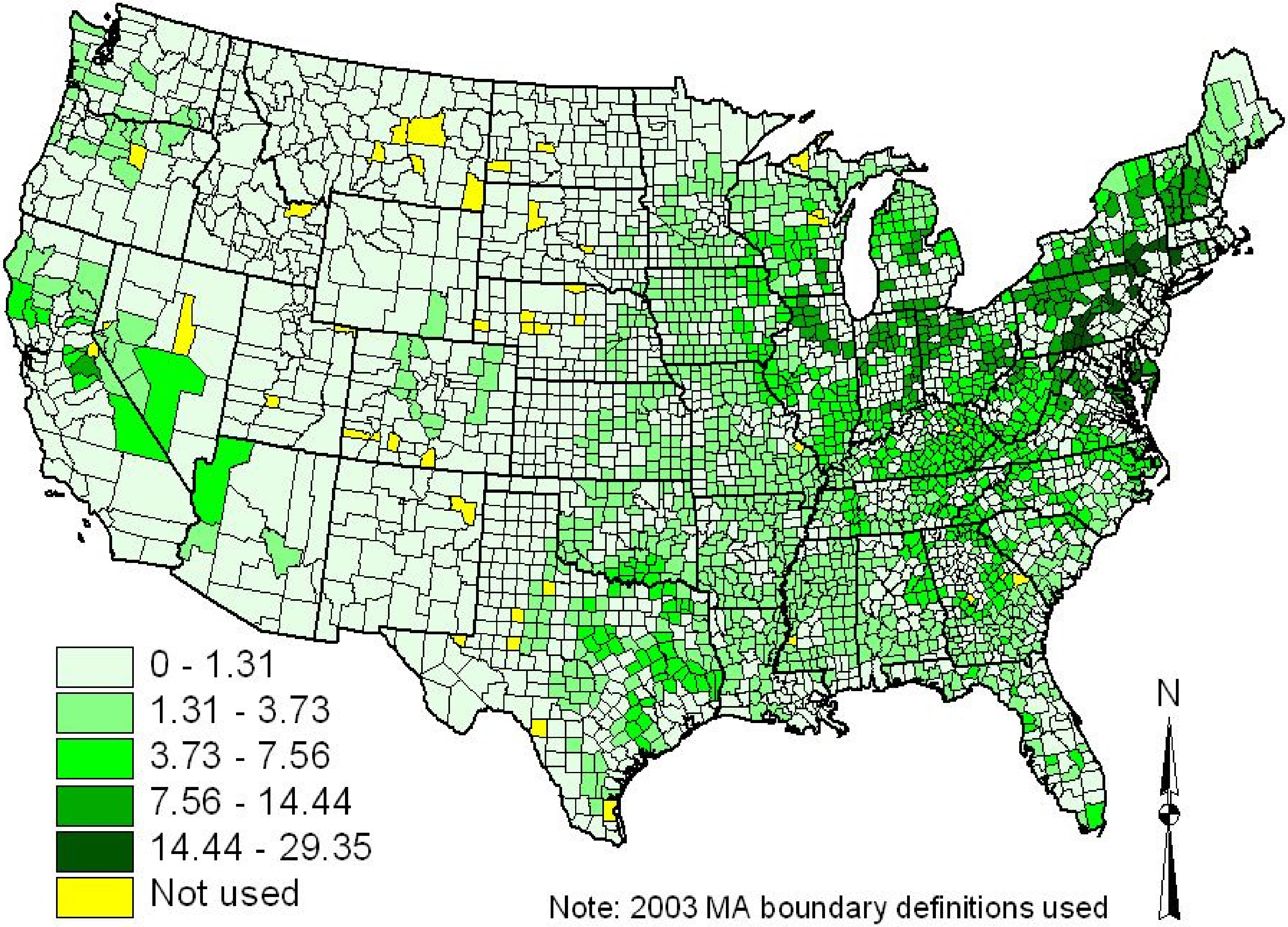
Figure 6: Distance Penalties (%) for Housing Costs 2000



Note: 2003 MA boundary definitions used

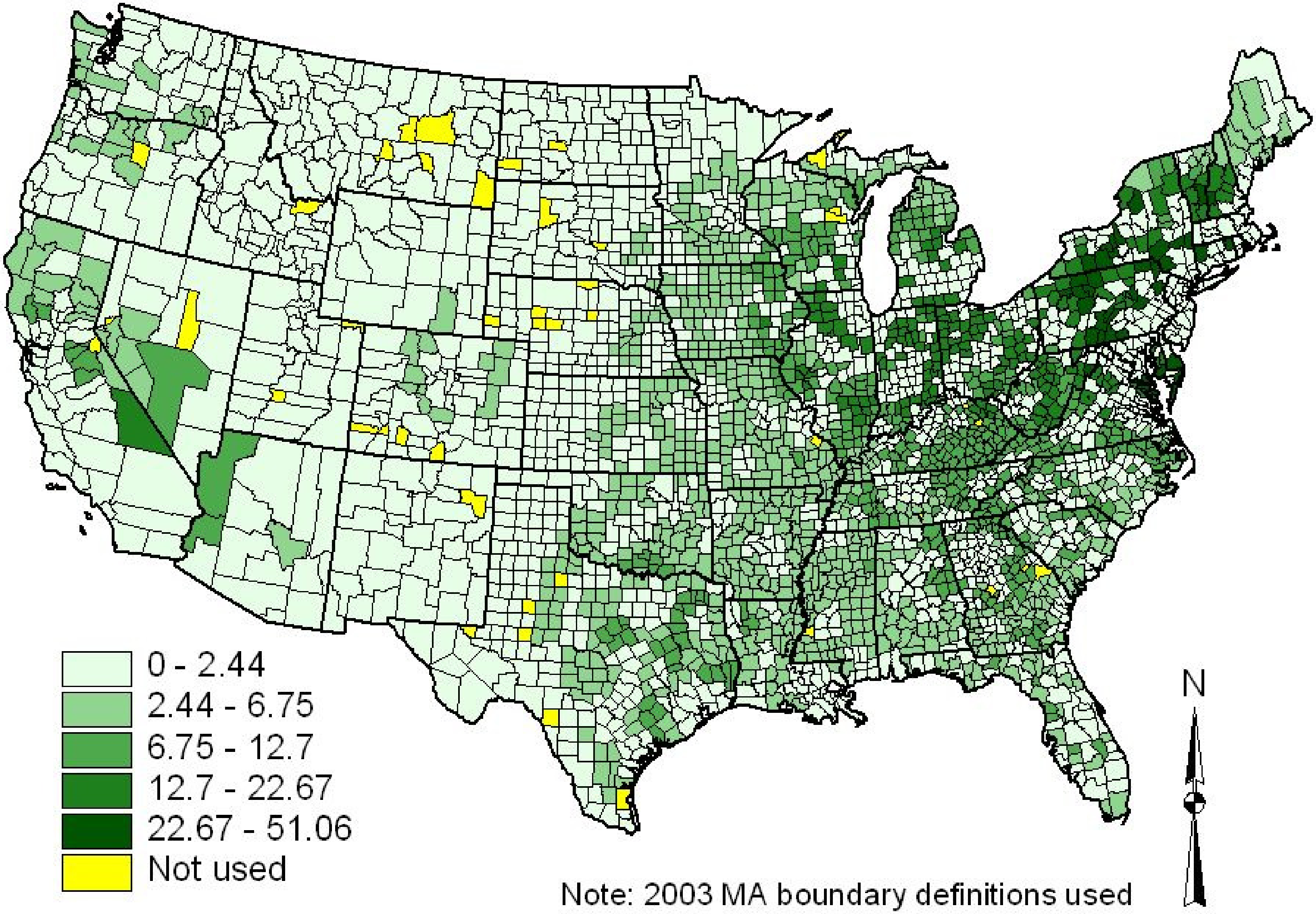
Note for reviewers: Color maps are in the pdf version.

Figure 7: Market Potential Effects (%) for Median Earnings 1999



Note for reviewers: Color maps are in the pdf version.

Figure 8: Market Potential Effects (%) for Housing Costs 2000



Note for reviewers: Color maps are in the pdf version.

Appendix Table 1. Variable Definitions and Descriptive Statistics (full sample)

Variable	**Description**	**Source**	**Mean**	**St. dev.**
Dependent Variables				
log(median earnings 1999)	Log of annual median earnings in 1999 (in dollars) for the employed 16+ pop.	2000 Census	9.89	0.18
log(wtd av med rent 2000)	Log of weighted average median gross house rent (dollars per month) of owner and renter occupied housing units 2000. For owner occupied units, imputed annual rent is calculated as 7.85% of median house value. Monthly average of that amount along with the median monthly rent for the renter occupied units are used to calculate the weighted average median rent, with weights being the shares of owner and renter occupied houses. Note that there are no official U.S. cost of living data series (in levels) at even the state level—let alone at the county level. Yet, research by the U.S. Department of Agriculture suggests that aside from housing cost differentials, other cost differences for non-housing goods and services tend to offset one another across urban and rural environments (Jolliffe, 2006a, 2006b).	2000 Census	6.18	0.37
Distance variables				
Dist to nearest/actual urban center (micropolitan or metropolitan area, CBSA)	Distance (in km) between centroid of a county and population weighted centroid of the nearest urban center, if the county is not in an urban center. It is the distance to the centroid of its own urban center if the county is a member of an urban center (in kms).	C-RERL	34.61	32.44
Inc dist to a metro	Incremental distance to the nearest/actual metropolitan area in kms (see text for details)	Authors' est.	36.68	49.06
Inc dist to metro>250k	Incremental distance to the nearest/actual metropolitan area with at least 250,000 population in 1990 in kms (see text for details)	Authors' est.	56.29	97.27
Inc dist to metro>500k	Incremental distance to the nearest/actual metropolitan area with at least 500,000 population in 1990 in kms (see text for details)	Authors' est.	40.67	66.83
Inc dist to metro>1500k	Incremental distance to the nearest/actual metropolitan area with at least 1,500,000 population in 1990 in kms (see text for details)	Authors' est.	89.77	111.47
Amenity/Ocean				
January Sun hours	Mean January sun hours	ERS, USDA	151.41	33.21
January temp	Mean January temperature (degree F)	ERS, USDA	32.95	12.07
July humidity	Mean July relative humidity (%)	ERS, USDA	56.15	14.49
July temp	Mean July temperature (degree F)	ERS, USDA	75.90	5.35
Topography Measure	A 1 to 24 score. 24 reflects the most mountainous terrain	ERS, USDA	8.83	6.59
Percent water	Percent of county area covered by water	ERS, USDA	4.61	11.29
Proximity to Great Lakes	1 if county centroid is within 50km of Great Lakes	Authors' est.	0.04	0.19
Proximity to Pacific Ocean	1 if county centroid is within 50km of Pacific Ocean	Authors' est.	0.02	0.13
Proximity to Atlantic Ocean	1 if county centroid is within 50km of Atlantic Ocean	Authors' est.	0.08	0.28
County area	County area in square miles	ERS, USDA	1011.13	1331.87
Market Potential Economic				
Agg hh inc within 100-200 km ring 1989 (mill.$)	Aggregate household income between 100 and 200 km radii from county centroid	1990 Census, Authors' est.	48687.36	58938.66
Agg hh inc within 200-300 km ring 1989 (mill.$)	Aggregate household income between 200 and 300 km radii from county centroid	1990 Census, Authors' est.	73681.00	74018.80
Agg hh inc within 300-400 km ring 1989 (mill.$)	Aggregate household income between 300 and 400 km radii from county centroid	1990 Census, Authors' est.	95508.52	83471.57
Agg hh inc within 400-500 km ring 1989 (mill.$)	Aggregate household income between 400 and 500 km radii from county centroid	1990 Census, Authors' est.	112480.84	92163.98
Population/Scale 1990				
County pop 1990	County population 1990	1990 Census	81806.94	268955.04
Pop of nearest/actual urban center 1990	1990 Population of the nearest/actual urban center measured as a micropolitan or metropolitan area	Authors' est.	375588.83	1381874.52
Labor Market Measures(Demography) 1990				
% African American 1990	% of 1990 population African-American	1990 Census	8.60	14.32
% Native American 1990	% of 1990 population that are Native American	1990 Census	1.44	5.59
% Hispanic 1990	% of 1990 population Hispanic	1990 Census	4.37	10.96
% Asian-Pacific 1990	% of 1990 pop Asian and Pacific islands origin	1990 Census	0.59	1.26

% Other ethnicity 1990	% of 1990 pop. with other race background	1990 Census	1.80	4.57
% 7-17 years 1990	% of 1990 population 7-17 years	1990 Census	16.78	2.34
% 18-24 years 1990	% of 1990 population 18-24 years	1990 Census	9.18	3.43
% 25-54 years 1990	% of 1990 population 25-54 years	1990 Census	39.74	3.71
% 55-59 years 1990	% of 1990 population 55-59 years	1990 Census	4.56	0.73
% 60-64 years 1990	% of 1990 population 60-64 years	1990 Census	4.70	0.98
% 65+ years 1990	% of 1990 population 65 years and over	1990 Census	14.97	4.33
% High school grad 1990	% of 1990 population 25 years and over that are high school graduates	1990 Census	34.36	6.12
% with some college 1990	% of 1990 population 25 years and over that have some college education	1990 Census	16.39	4.50
% with associate degree 1990	% of 1990 population 25 years and over that have an associate degree	1990 Census	5.34	2.10
% College grad 1990	% of 1990 population 25 years and over that are 4-year college graduates	1990 Census	13.43	6.45
% Female 1990	% of 1990 population that are female	1990 Census	51.02	1.61
% Married 1990	% of 1990 population that are married	1990 Census	59.12	6.23
% with disabilities 1990	% of 1990 16-64 pop with a work disability	1990 Census	9.55	2.89
Labor Market Measures (Demography) 2000				
% African American 2000	% of 2000 population African-American	2000 Census	8.75	14.46
% Native American 2000	% of 2000 population that are Native American	2000 Census	1.53	5.85
% Hispanic 2000	% of 2000 population Hispanic	2000 Census	6.13	12.04
% Asian-Pacific 2000	% of 2000 pop Asian and Pacific islands origin	2000 Census	0.81	1.59
% Other ethnicity 2000	% of 2000 pop. with other race background	2000 Census	2.57	4.86
% 7-17 years 2000	% of 2000 population 7-17 years	2000 Census	16.50	2.03
% 18-24 years 2000	% of 2000 population 18-24 years	2000 Census	8.86	3.32
% 25-54 years 2000	% of 2000 population 25-54 years	2000 Census	41.12	3.51
% 55-59 years 2000	% of 2000 population 55-59 years	2000 Census	5.21	0.84
% 60-64 years 2000	% of 2000 population 60-64 years	2000 Census	4.50	0.96
% 65+ years 2000	% of 2000 population 65 years and over	2000 Census	14.82	4.10
% High school grad 2000	% of 2000 population 25 years and over that are high school graduates	2000 Census	34.79	6.52
% with some college 2000	% of 2000 population 25 years and over that have some college education	2000 Census	20.40	4.34
% with associate degree 2000	% of 2000 population 25 years and over that have an associate degree	2000 Census	5.70	1.98
% College grad 2000	% of 2000 population 25 years and over that are 4-year college graduates	2000 Census	16.43	7.67
% Female 2000	% of 2000 population that are female	2000 Census	50.50	1.90
% Married 2000	% of 2000 population that are married	2000 Census	57.70	5.41
% with disabilities 2000	% of 2000 16-64 pop with a work disability	2000 Census	11.95	3.20
House characteristics 1990				
House age 1990	Age of housing unit in 1990 (years)	1990 Census	26.04	9.57
Share 1 bedroom 1990	Share of 1 bedroom house to total rooms 1990	1990 Census	0.09	0.04
Share 2 bedroom 1990	Share of 2 bedroom house to total rooms 1990	1990 Census	0.32	0.05
Share 3 bedroom 1990	Share of 3 bedroom house to total rooms 1990	1990 Census	0.43	0.06
Share 4 bedroom 1990	Share of 4 bedroom house to total rooms 1990	1990 Census	0.12	0.04
Share 5 bedroom 1990	Share of 5 bedroom house to total rooms 1990	1990 Census	0.03	0.02
Share mobile homes 1990	Share of mobile units to all housing units 1990	1990 Census	0.14	0.08
Share complete plumb 1990	Share with complete plumbing facility 1990	1990 Census	0.97	0.03
Share complete kitchen 1990	Share with complete kitchen facility 1990	1990 Census	0.98	0.02
House characteristics 2000				
Median rooms 2000	Median number of rooms 2000	2000 Census	5.44	0.40
House age 2000	Age of housing unit in 2000 (years)	2000 Census	31.40	10.62
Share 1 bedroom 2000	Share of 1 bedroom house to total rooms 2000	2000 Census	0.09	0.04
Share 2 bedroom 2000	Share of 2 bedroom house to total rooms 2000	2000 Census	0.30	0.05
Share 3 bedroom 2000	Share of 3 bedroom house to total rooms 2000	2000 Census	0.43	0.07
Share 4 bedroom 2000	Share of 4 bedroom house to total rooms 2000	2000 Census	0.13	0.05
Share 5 bedroom 2000	Share of 5 bedroom house to total rooms 2000	2000 Census	0.03	0.02
Share mobile homes 2000	Share of mobile units to all housing units 2000	2000 Census	0.15	0.10
Share complete plumb 2000	Share with complete plumbing facility 2000	2000 Census	0.98	0.02
Share complete kitchen 2000	Share with complete kitchen facility 2000	2000 Census	0.98	0.02
Number of counties			**3028**	

Notes: The metropolitan/micropolitan definitions follow from the 2003 definitions. ERS, USDA = Economic Research Services, U.S. Department of Agriculture; C-RERL = Canada Rural Economy Research Lab, University of Saskatchewan.

Appendix Table 2: Mean and Standard Deviations (in parentheses) of Major Variables by Population Group

Variables	Non-metro	Rural	Small MA≤ 250k	Large MA>250k
Dependent variables:				
log(median earnings 1999 in $)	9.82	9.80	9.93	10.08
	(0.14)	(0.13)	(0.14)	(0.17)
log(wtd av med rent 2000 in $/month)	6.04	5.97	6.31	6.51
	(0.32)	(0.33)	(0.26)	(0.32)
Dist to the nearest urban center	41.07	59.91	n.a.	n.a.
	(36.52)	(30.56)		
Dist to the center of own metro	n.a.	n.a.	17.76	28.60
			(18.61)	(19.52)
Incremental distance to a MA	55.40	43.47	n.a.	n.a.
	(51.67)	(49.93)		
Incremental distance to MA>250,000	66.80	76.02	93.23	n.a.
	(106.20)	(115.19)	(93.26)	
Incremental distance to MA>500,000	42.89	45.32	36.89	36.29
	(66.07)	(68.95)	(59.07)	(73.34)
Incremental distance to MA>1,500,000	89.03	83.45	78.54	99.37
	(111.10)	(106.24)	(115.44)	(139.88)
No. of observations	1972	1300	416	640

Notes: The categories are determined using 2003 micropolitan and metropolitan area definitions. See the text for more details.