Population Density across the City: The Case of 1900 Manhattan*

Jason Barr† Teddy Ort

March 2014

Rutgers University, Newark Working Paper #WP2014-005

Abstract

The literature on urban spatial structure tends to focus on the distribution of residents within metropolitan areas. Little work, however, has explored population density patterns within the city. This paper focuses on a particularly important time and place in urban economic history: 1900 Manhattan. We investigate the determinants of residential spatial structure during a period of rapid population growth. We explore the effects of environmental factors (such as historic marshes and elevation), immigration patterns, the location to amenities, and employment centers. While environmental features and amenities were significant, their effects were dominated by the location choices of immigrants. On the supply side, we also explore how the grid plan affected density.

key words: population density, Manhattan, spatial structure, immigration

JEL Classification: N3, N9, R2

---

*We thank Eric Sanderson for sharing his data on Manhattan's historical ecology. We thank Brian Engelmann for his help with creating the shapefile for 1900 Manhattan. We appreciate Froozan Makhdoom's help with data collection. We also appreciate comments from Jan Brueckner, Gergely Baics, Troy Tassier, Francesc Ortega, Michael Edelstein, and the seminar participants at Fordham University and CUNY Queens. All errors belong to the authors.

†Corresponding author. Department of Economics, Rutgers University Newark, Newark NJ 07012. email: jmbarr@rutgers.edu.

1 Introduction

During the late-19th century, rapid industrialization and massive immigration – both foreign and domestic – caused city populations to swell. Land annexations meant that cities were able to preserve the traditional urban density of 125 to 150 people per acre (Condit, 1980). Nonetheless there was tremendous variation within these cities.

While Manhattan's density in the 1890s was around 143 people per acre, some neighborhoods had densities of over 800 people per acre; these blocks were considered the densest spots on planet earth. Paris' density was around 125 people per acre, but the Arrondissement Bonne Nouvelle had a density of 3.5 times that. The Koombarwara neighborhood in Mumbai had some 760 people per acre.[1] Table 1 shows the density of selected cities and particular neighborhoods.

City	Density	Year	Neigborhood	Density	Year
New York*	143.2	1893	Eleventh Ward**	964.4	1894
Paris	125.2	1891	Bonne Nouvelle	434.2	1891
Glasgow	92.5	1891	St. Rollox	350.0	1891
Mumbai	57.7	1893	Koombarwara	759.7	1881
London	57.7	1893	Bethnal Green, North	365.3	1891

Table 1: Population densities of selected cities, and particular neighborhoods in the late-19th century. Note neighborhood names are taken directly from the table. Source: Tenement Housing Commission (1895), pages 256-257. *Manhattan only; exludes annexed sections of the Bronx. **Sanitation district A, within the 11th ward.

Despite the rapid growth of urban populations after the Civil War, little work has explored how these population bursts determined land use patterns within the city. Our work aims to understand the forces that drive intra-urban residential land use, by looking at the factors that caused people to cluster at particular locations and avoid others. By focusing on a case in urban history, we can see how the spatial structure formed in a period with relatively few building and zoning regulations, and before the automobile. The decisions made about land use then continue to have ramifications today, given the durable nature of real estate, historical lock-in due to economies of scale and zoning regulations, which tend to "freeze" land use types when implemented.[2]

The literature on spatial structure tends to focus on two variables in determining population density: distance to the center – as a proxy for commuting time and costs – and income. The general finding is that lower income families cluster near the city center in dense concentrations to reduce their commuting costs; higher income families

[1] In 1890, Chicago's population density was 18.79 people per acre (Report of Vital and Social Statistics, 1890, Volume 2, page 3). In 1900, the highest density neighborhood had 55 people per acre (Parsons, 1916). Boston's 1890 population density was 19.59. (Report of Vital and Social Statistics, 1890, volume 2, page 3).

[2] We estimate that about 24% of Manhattan's developed land (including parks) contains a structure completed before 1916, when the city's first zoning rules were implemented (data from NYC Deparment of City Planning, 2013).

live further from the center so they can consume more housing at a lower price (Muth, 1969).[3] While these results strongly predict density patterns within metropolitan regions, they are limited with respect to understanding the spatial structure within a city.

Beyond commuting costs and income, other factors drive the location decisions of households. For example, all else being equal, people prefer to live with others who share their cultural backgrounds, such as immigrants who prefer to live with their compatriots (Hatton and Williamson, 1994). Cities also have a wide range of local amenities, such as parks, entertainment, and shopping venues; access to these will influence household residential decisions.

In addition, the natural environment may also play a role in shaping residential patterns. Historically, cities were often developed by covering rivers, draining marshlands, and flattening hills. Since former marshlands used for residences were sometimes filled in with garbage and waste, and these neighborhoods were prone to flooding, these areas likely saw an increase in the prevalence of disease. On the other hand, higher elevations were seen as more salubrious, since the ground remained dry during wet seasons (Citizens' Association, 1865). In short, the demand for any particular location will be influenced by not only the distance to the business districts, but also the personal, social, and environmental benefits that a particular location can provide.

On the supply side, platting schemes can also determine land use patterns, because the size and location of available plots will affect the costs and types of housing provided. Little work has explored how land divisions have affected the provision of housing and population density. In this paper we provide some evidence on the effects of the 1811 grid plan, which imposed relatively large blocks north of Houston Street, while leaving intact the smaller, pre-1811 blocks in the southern part of the city. At the time, reformers argued that increased plot sizes would increase the size and efficiency of land use and decrease overall density (Flagg, 1884; Veiller, 1903a). We find that larger blocks were associated with lower density, even after controlling for immigrant patterns, and the distance to City Hall. This provides evidence that platting schemes can affect spatial structure.

Turn of the 20th century Manhattan provides a useful case study. First, it drew millions of immigrants, foreign and domestic, who came to the city to find work. Manhattan Island was the City of New York until 1874 when it annexed the western part of the Bronx (and the rest of the borough in 1895). In 1898, the city annexed the City of Brooklyn, Queens and Staten Island. While Brooklyn and the Bronx had a few tenement neighborhoods, they were neither as extreme nor as numerous. In 1900, Manhattan contained 54% of New York's population and 52% of its tenements, yet only 8% of its total land mass (Veiller, 1903b).

Second, Manhattan is a narrow and long island, which can constrain the competitive process. The island is about 13 miles long, but only 2.5 miles across at the widest point. The average distance between the Hudson and East rivers is only 1.5 miles. In lower Manhattan, just north of City Hall and the central business district is the Lower East

[3] Although European cities have negative exponential density functions (Anas, et al., 1998), the income-based spatial structure appears, in several cases, to be different than that of U.S. cities. As discussed below, Brueckner et al. (1999) model how this pattern may be determined by the location of amenities.

Side. This neighborhood is particularly interesting because in the late 19th and early 20th centuries it was considered the densest neighborhood on planet earth. To our knowledge, no work has explored why such extreme density emerged at this particular location.

To investigate the role of neighborhood characteristics on population density, we have collected a new and unique data set for 1900 Manhattan. Volume II of the First Report of the newly-formed Tenement House Department of the City of New York (1903) includes a set of maps and tables containing data, at the block level, for the vast majority of blocks on Manhattan Island. Data include population, density, average household size, native or foreign-born status, and the birth country of the parents of household heads. The data are block-level amalgamations from the 1900 census.

As controls, we have collected a number of important variables that allow us to test how topographical, environmental, and location- and amenity-related factors drove the spatial distribution of the population within the island. Such variables include the distances to parks, houses of worship, and elevated railroad stops; and the pre-European natural environment, such as the presence of marshlands and rivers, and the elevation. For platting, we look at the average size of blocks formed by the grid plan of 1811. Presumably, large blocks would make assemblages easier, and reduce structural density. We also investigate how immigration patterns affected the density distribution. Since immigrant neighborhood choice and population density are likely to be endogenous, we employ two-stage and three-stage least squares methods, using past immigration and demographics as instrumental variables.

The three-stage regressions allow us to offer some evidence on the preferences of the immigrants themselves. We find that immigration location is not only determined by time in the country (and income), and the desire to cluster with fellow countrymen, but preferences about other immigrant groups. Because some nationalities can view other nationalities either favorably, disfavorablty or neutrally, this can affect where and how dense settlement occurs. For example, if Catholics and Protestants view living near each other as problematic, it can reduce overall density in some neighborhoods, if this rivalry causes the population to spread out more than would be the case otherwise. A similar effect can occur if Christians and Jews view each other negatively.

To summarize our results, we find that, first, distance to center is less relevant at the local level, once additional controls are included. There is a strong interaction with historical density patterns and the distance to the center variable, which provides evidence that distance to the center is a proxy for commuting costs for low-wage families, who need to be near their jobs, to save on expenses, and who generally can only afford the lower-quality housing stock found in these locations. Second, the results support that earlier environmental and ecological variables continued to influence density patterns at the turn of the 20th century–suggesting both a lock-in effect from earlier periods, and the ability of even low-income groups to avoid the worst neighborhoods.

Third, in regard to amenities, we find that densities were lower for blocks that bordered parks, suggesting the rich outbid the poor there, but beyond that there was a negative relationship between park locations and distance. As well, density fell off rather steeply with distance to elevated railroad stops. Fourth, we find a strong relationship between immigration patterns and density. Groups that came earlier (the Germans and the Irish) tended to be more spread out on the island and lived in less dense neighborhoods. Newer immigrant groups (such as the Russian and Polish Jews, and the Italians), lived closer to downtown and in more dense neighborhoods. In addition we find that the Germans were considered a type of neighborhood amenity. All the other major immigrant groups were attracted to neighborhoods with large German concentrations. On the other hand immigrant groups preferred to avoid the other major groups. Next we turn to a specific focus on the Lower East Side, the most densely populated neighborhood on planet earth in 1900. We find that the variation in density across blocks in this area can be explained by quality and quantity of housing stock measures, as well as the distance to synagogues. Lastly, on the supply side, we find that block sizes imposed by the grid plan influenced density, with pre-grid smaller blocks having larger densities.

The remainder of this paper is organized as follows. The next section discusses the relevant literature. After that, in Section 3, we the theory land use within a spatial equilibrium context in order to generate testable hypotheses. Then Section 4 discusses the data set and the variables we investigate. Next, in section 5, we present the results of different sets of regressions and tests. Finally, we offer some concluding remarks in Section 6. Two appendices provides more details on the data, sources and preparation.

2 Relevant Literature

This paper is related to several different literatures. The first area relates to urban spatial structure. A large body of research has focused on the relationship between urban density and distance from the urban core (McDonald, 1989; Anas, et al., 1998). Recent modeling has also aimed to understand metropolitan area poly-centric development (Fujita and Ogawa, 1982; Helsley and Sullivan, 1991) and sprawl (Glaeser and Khan, 2004). However, there is relatively little work on understanding the patterns of intra-urban spatial structure, specifically on the historical causes of extreme density. In recent work, Glaeser, et al. (2008) conclude that access to public transportation can account for the spatial segregation between rich and poor. Lower-income households move to central cities to be close to public transportation, while upper income households tend to travel by cars and thus choose to live outside the central cities where they can buy more land at lower prices.

Our work is similar in spirit to Muth (1969), who models the causes of population density across the city. He tests the model using data from Chicago within the Southside neighborhood. We present a model of density similar to Muth. However, our empirical work, while building on his, includes a more detailed treatment of the spatial factors that drive urban density. In addition, our model of density is similar to Brueckner et al. (1999), who show how urban amenities may affect the urban rent gradient and the distribution of rich and poor within an urban area.

Another branch of research investigates the effects of natural advantage or environmental factors on the location of economic activity. For example, Davis and Weinstein (2002) test how natural advantage may have locked-in metropolitan area population levels in Japan. Bleakley and Lin (2012) find that the location of historic portages predicts density today. In this sense, regions that were appealing during initial settlement can produce increasing returns. These persist because of infrastructure investments and economic development that produce spillovers and economies of scale, which continue to remain important even after the initial advantages are not.

Regarding "negative" local effects, our work is related to Villarreal (2013), who investigates how historic marshlands have determined housing rent values, finding negative effects that persisted throughout the 20th century. Burchfield, et. al. (2006) find that geology and topography can create natural barriers to sprawl. Saiz (2010) finds that cities surrounded by more bodies of water have lower housing price elasticities with respect to land. However, Barr, Tassier and Trendafilov (2010) test for the effect of bedrock depth on the spatial distribution of skyscrapers and find a very small effect. Our work is similar in that we test how the presence of marshlands, rivers, and topography have influenced historical residential patterns, and whether these factors persisted until 1900. We find that some earlier environmental features where important in affecting the later density patterns of the city.

Other historical work on New York City includes Margo (1996), who investigates apartment rent values in New York from 1830 to 1860. He finds a significant rent gradient going north from the central business district. This finding supports that accessibility, especially for working class households, is crucial. Atack and Margo (1998) explore historical land values in Manhattan. They find that prices rose quite dramatically over the 19th century, as population and economic development increased. Further, they find a very steep land value gradient before the Civil war, and a flattening of the gradient during the second half of the 19th century. Transportation developments during the century allowed for a northward movement of economic activity. We find a diminished value for the distance to City Hall variable, once we account for the spatial distribution of immigrant groups, which suggests that the longer a group remained in the city, the more decentralized was its distribution.

Existing work on immigration includes Hatton and Williamson (1994) who demonstrate that early emigration patterns are consistent with the friends and neighbors theory, i.e., new immigrants settle in ethnic enclaves. Work on immigration and ethnic segregation includes Cutler, et al. (2008), who find segregation is largely driven by cultural difference between immigrants and natives; immigrants prefer to live with residents from their home countries. Here we investigate how the presence of different immigrant groups affected the location choices of other groups. Some enthnicities acted as "attractors" and others as "repellents." We provide some evidence on how immigrant preferences affected spatial structure.

Finally, this work revisits a large historical literature on tenements and slums. While earlier writings focused on tenement conditions and health issues, little work has attempted to understand the economic determinants of such densities. One of the few economic works from that time is that of Pratt (1911) who determined that population density in lower Manhattan was driven primarily by the distance to factories. Much of the literature of that time was written to expose the plight of the poor regarding their sanitary, health and living conditions (Citizens' Association, 1866; Riis, 1890; DeForest and Veiller, 1903).

3 The Theory of Intra-Urban Density

We are interested in population density at the city block level.[4] Density for block i is given by the population per square acre of land: $\Phi_i = \Psi_i / L_i$, where Ψ_i is the population of the block and L_i is the land area. Population density can also be defined by the ratio of structural density to housing consumption per household times average family size (Muth, 1969; Brueckner, 1987) :

$$\Phi_i = \left(\frac{\Psi_i}{L_i}\right) = \left(\frac{\Psi_i}{N_i}\right)\left(\frac{H_i}{L_i}\right) / \left(\frac{H_i}{N_i}\right) \equiv F_i \left(\frac{S_i}{D_i}\right),$$

where N_i is the number of households on the block. We define the amount of housing consumed per household as $D_i \equiv H_i / N_i$, which is a measure of the demand for housing per household. $S_i \equiv H_i / L_i$ is the quantity of housing supplied per unit of land (i.e., structural density). Finally, $F_i \equiv \Psi_i / N_i$ is the average family or household size for the block. Denoting lower case letters as the log of a variable gives $\phi_i = f_i + s_i - d_i$.

Note that anything that increases the demand for housing per household will reduce the density. That is, if a household prefers a larger home, all else equal, it will consume more space per person and hence there will be fewer people per acre. On the other hand, an increase in the amount of housing per acre, ceteris paribus, will increase density. If, for example, a given lot has a building of four floors, and an adjacent lot of the same size has a building with five floors, there will be an additional group of people living in the taller building, which results in a higher density. Note that density will be positively related to housing prices. On the supply side, a higher price will increase the structural density (quantity supplied) and reduce the quantity of housing demanded per household. We now turn to the determinants of the demand for housing. Note that for the sake of brevity we do not include a model of housing supply (for more information on the determinants of structural density, see Brueckner 1987)

3.1 Demand for Housing

If we assume that households are identical, then the demand for housing by a household at a particular location emerges from the maximization of a utility function over housing, H, and a numeraire Z , $U(Z, H)$, subject to a budget constraint, $Y - TX = Z + PH$, where Y is household income, $T > 0$ is the cost of commuting per mile to work, assumed fixed across locations, and X is the distance of a residence to the center.[5] TX is the total commuting cost, and $Y - TX$ is net income. Z is assumed to have a fixed price of one (Muth, 1969; Brueckner, 1989).

[4] We do not distinguish between net density and gross density. Rather, we assume gross density at the block level, given that it is our dependent variable in the empirical sections. The problem of net versus gross density is somewhat diminished at the block level compared to other levels of aggregation such as census tract or zip code level. However we return to the issue of net density in section 5.4.

[5] Note that for simplicity, in this subsection we denote H as the amount of housing consumed per household, where above H is the total housing consumed in a block. We do this to reduce the number of variable designations in the model, without, we hope, sacrificing clarity.

The first order condition gives $U_2/U_1 = P(X)$, where U_i is the partial derivative of the $i^{th} = 1, 2$ variable. Thus, a quantity of housing is chosen such that the marginal rate of substitution is equal to the relative price of housing. If we assume an open city, where immigration and population movement creates a spatial equilibrium with equal utility across households, denoted μ, then the price of housing at a location X, will be set so that

$$U(Y - TX - P(X)H, H) = \mu. \tag{1}$$

Taking the differential of equation (1) with respect to X, yields the standard spatial equilibrium result:

$$-HP_X = T. \tag{2}$$

This condition says that if a household moves a small distance further from the center, the marginal expenditures on housing will be reduced by an amount equal to the increased transportation costs. Thus, there can be no net gain of utility across locations when all households are in equilibrium. The price of housing must drop with distance to the center (i.e., $P_X = -T/H < 0$) to compensate the household for the increased cost of transportation. Since prices are lower further from the center this will both increase housing consumption per household, and decrease structural density, yielding a drop in population density moving away from the center.[6]

3.2 The Role of Income

If a household should experience a change income, it will possibly change its preferred location. Totally differentiating equation (2) yields

$$P_X dH + H dP_X + dT = 0. \tag{3}$$

This condition says that in equilibrium, three factors that must be "balanced" when there is some exogenous change (in this case income). That is, any change in housing consumption, prices and/or transportation costs must be offset, so no net gain in utility is possible.

First, if income changes, it will increase the opportunity cost of commuting, $dT = T_Y dY > 0$. Second, a rise in income changes the the desired amount of housing; but since housing is not independent of location, we have the condition: $dH = H_Y dY + H_p^c P_X dX$, where $H_Y dY > 0$ is the increased demand for from more income (given that housing is assumed to be a normal good) holding location constant, and $H_p^c P_X dX > 0$ is the (income compensated) change in housing that comes from a move to a new location. Finally, movement affects the price gradient, which is assumed to flatten further from the center, i.e., $dP_X = P_{XX} dX > 0$ (Muth, 1969).

Equation (3) can be re-written

$$\left[(P_X)^2 H_p^c + P_{XX}\right] dX + \left[P_X H_Y + T_Y\right] dY = 0. \tag{4}$$

[6] If we define $\rho = T/PH$ as the transportation cost per mile relative to housing expenditures then equation (2) can be written $dP/P = -\rho dX$, which generates the negative exponential price gradient, i.e., $P = C_0 e^{-\rho X}$.

The assumption is that $P_X H_Y + T_Y < 0$,which means that movement away will produce a net benefit since the gain from additional housing expenditures offsets the increased opportunity cost of commuting. How far away it moves depends on $\left[(P_X)^2 H_p^c + HP_{XX}\right]$,which, by assumption, is positive. It measures the steepness of price drops from the center, and is assumed to level off moving away (Muth, 1969). Solving equation (4) shows that $dX/dY > 0$.Thus, if a household experiences a rise in income, it will consume more housing in the outer ring of the city, which means a drop in population density in these areas. In short, the model predicts that prices will be higher closer to the center because of lower transportation costs, and further that higher income people will move to the suburbs, as long as housing is a normal good and transportation costs don't rise too quickly with income. Since the price of housing relative to income will be much higher closer to the city it will further bid up density, as the poor reduce their per-person consumption, and hence will see over-crowded tenement districts when population growth is rapid.

3.3 The Effect of Amenities

The spatial distribution of the population can depend on the distribution of amenities or disamenities, and the value that different residents place on these amenities.[7] A valuable amenity will raise the value of a particular location, and thus there will be higher housing prices, which will tend to increase density, cet. par. A disamenity, such as housing on former swampland, will reduce the price of housing and reduce density. However, the net effect on density depends on the strength of preferences among rich and poor. If the wealthy have a high value for a particular location, they can outbid the poor to live at that location. The net effect, will be to reduce population density to a degree, since the wealthy will use their additional income for more housing.

Let's say there is an exogenous amenity in the neighborhood so that utility is given by

$$U\left(Y - TX - PH, H; A\right)$$

where $A > 0$ is the quantity of the amenity (Brueckner et al., 1999). We assume that it doesn't directly affect the quantity of housing, though it will indirectly influence it via the price. Assume, for example, that there's an amenity in the center that dissipates moving away from the center, so that $A_X < 0$, $A_{XX} > 0$. The spatial equilibrium condition is

$$-P_X H = T - P_A A_X \tag{5}$$

or $P_X = -\frac{1}{H}\left[T + |P_A A_X|\right] < 0$.[8] $P_A A_X < 0$ is the reduced value (utility cost) of the amenity when moving further away from the center. Equation (5) says that in a spatial

[7] Note that we just focus here on the role of a general amenity. A larger, more extentive model could include the role of immigration, and how immigrant prefences could interact to affect density patterns across the city and the price of housing. We leave this expansion for future work.

[8] Note that in equilibrium, $U_3/U_1 = \lambda$; where, λ is the "price" of A. Since A is the quantity of the amenity, its cost will be the increase in the price of housing due to the amenity. The price of A is the change in housing price from a small increment in the amenity. This result can also be shown using the first order conditions of the Lagrangian function.

equilibrium any change in marginal housing expenditures with a small movement away from the center must now be offset by both the transportation costs and the lost amenity value. Note that housing prices further out will drop even more compared to the case with no amenities since there are two forces driving price changes, transportation costs and less amenity. Closer to the amenity will have greater the population density, because the higher the price of space.

The net effect of a the location choice of a household with an rise in income now depends on how it values the amenity relative to more housing. Totally differentiating equation (5) yields

$$P_X dH + HP_{XX} dX - (P_A A_{XX} + P_{AX} A_X)\, dX = 0. \tag{6}$$

For simplicity, and without loss of generality, we assume that $T_Y = T_X = 0$. Furthermore,

$$dH = H_Y dY + H_p^c \left[P_X + P_A A_X\right] dX, \tag{7}$$

which says that changes in housing consumption can come from a rise in income or movement away from the center. However, movement has two effects. The first effect is $H_p^c P_X > 0$, i.e., movement lowers the price of housing and increases the housing quantity demanded. The second effect is related to the amenity. Movement away will reduce the amenity but will also reduce the price of housing, so $P_A A_X < 0$, and thus $H_p^c P_A A_X > 0$. That is, movement away will increase the quantity of housing demanded due to less amenity, i.e., the amenity and housing are substitutes. In addition, similar to above, we assume $P_{XX} > 0$, $A_{XX} > 0$, and also $P_{AX} < 0$. Solving for the change in distance with respect to income now gives

$$dX = -\frac{P_X H_Y}{\left(P_X^2 H_P + HP_{XX}\right) - \left(P_A A_{XX} + A_X P_{AX}\right) + \left(H_P P_X\right)\left(P_A A_X\right)} dY \equiv \beta_1 dY, \tag{8}$$

When there is an amenity, there are now additional terms in the denominator. $(P_X^2 H_P + HP_{XX}) - (P_A A_{XX} + A_X P_{AX})$ is the derivative of $(P_X H - P_A A_X)$, and, once again, is assumed to be positive. The other term is $(H_P^c P_X)(P_A A_X) < 0$, since housing and the amenity are substitutes, movement away from the center will reduce the value of the amenity in proportion to the increase of housing. If this loss is large, it will result in a higher-income family moving closer to the city, i.e., the family gives up housing in exchange for more of the amenity. As a result, wealthier families can outbid the lower income families for the choice locations near the center.[9]

β_1 is, roughly speaking, how far out a household will move with a \$1 rise in income. The location choice with the presence of an amenity is not necessarily positive because the denominator now contains both negative and positive terms (assuming it's not exactly zero). If the first term dominates the second, the denominator will be positive and movement away will occur (although, not as far as if there were no amenities). If the second term dominates then more affluent households will move closer to the center because

[9] This is similar in spirit to Brueckner et al. (1999) who look at bid rent functions for land, and show that higher income families can outbid lower income families in order to be close to the amenity.

the utility value of the amenity (and its changes) are so great that the increased housing consumption available further out will not be enough to compensate the households.

In short, a rise of income can influence the incentive to move away from the center in a positive or negative direction. By moving, those with higher income can consume more housing at a lower price, but they are farther away from the amenity. If the utility value of the amenity is very high and the utility value drops off quite steeply with distance, then the increase in housing gained by moving might not offset the loss of being farther away from the amenity.

While some amenities, like shopping, may not yield a steep utility loss, other amenities may be strongly tied to the center. In particular, people prefer to be close to their social and demographic groups and moving to a different neighborhood could conceivably produce a large utility loss in this respect. Similarly, with parks, we can imagine that parks provide views, cleaner air, more light, recreation, and less density by removing parts of the city from development. Moving away from a park will reduce these benefits rather quickly and could provide a steep drop in utility that is not compensated by more housing. Notice, too, that the magnitude of β_1 depends on the forms of the underlying functions. If one of those forms should change, β_1 will change and the equilibrium distance will change as well.

In short, housing consumption per household follows income. If wealthy people find that more housing and less amenity is utility maximizing, we expect to see less density further out. If the wealthy prefer to live closer to the center, they will consume more housing there, and thus density will increase moving away from the core.

3.4 The Theory and Manhattan

The model predicts that where there are important amenities in the city that are highly desirable to upper income households, we would see them outbidding the lower income households to be closer to the amenities. Although the wealthy pay a higher price for housing, their higher income means they can still consume more housing and have lower population density (though at greater densities than without the amenity). However, this high price will encourage high structural density. This can be called the "Central Park Effect." All along Central Park East and West we see block after block of large, high-rise apartment houses for the wealthy; the length of the park is some 2.5 miles on both sides. Living adjacent to the park affords great views, reduces congestion, and, of course, allows easy access to the park itself. As such, many wealthy families outbid lower income ones for the right to live next to the park. A similar pattern of land use can be seen along Riverside Park adjacent to the Hudson River.

Manhattan also contains some historic disamenities. In particular, lower Manhattan is very close to sea level; a substantial fraction of the area contained marshlands, rivers, and other bodies of water such as Collect Pond. These wet areas were likely to promote water-borne or insect-carried diseases, such as cholera and yellow-fever, since they had poor drainage. However, lower Manhattan was also the source of employment for many lower-income and immigrant households. As such, these people had to weigh the negative aspects of living near or on top of a disamenity versus the benefits of being close to the center. If they perceived that the likelihood of getting deathly sick from living downtown was high, they would lower their demand for housing there and therefore density would decrease. On the other hand, if being close to the center had a high return relative to the probability of losing income, we would expect to see no drop in the density of these areas. We would, overall, expect that a disamenity in a neighborhood would lower the price of housing in that area for all income groups and result in lower density.

Uptown, however, where the land was dryer and less swampy, density could increase or decrease depending on the interaction between income and preferences. If the wealthy outbid the lower income households for the highest elevations, then we would expect to see a negative relationship between income and density. Ultimately, when looking at an amenity and density, it remains an empirical question about who values the amenity more (or values being away from a disamenity) – lower or upper income groups. When the lower income group values something more there will be a positive relationship between density and amenities; when the upper income households do, there will be a negative relationship. Regardless, the more attractive a neighborhood, and the greater the price of housing, the more structural density there will be, but the demand by residents will determine if the property contains a luxury apartment buildings such as the Dakota, or a row of tenements.

Lastly, immigration is likely to affect land use based on two factors: the income levels of particular groups and their desire to live in ethnic enclaves. Higher incomes is likely to cause immigrant groups to move away from the center, decreasing density. Yet, if they desire to cluster together–which can be seen as an neighborhood amenity–they may be drawn to particular areas of the city, increasing density.

4 The Data

Generally speaking, we have three categories of data: demographics, amenities and locational. We review each in turn. Table 2 gives the descriptive statistics. Appendix A contains more details.

4.1 Demographics and Immigration Data

Demographic data include the population density of each block, the average household size (total block population divided by number of households), and immigrant details.[10] The immigrant variables we work with are the number of heads of households (HoHs) with parents from a particular country. Figure 1 shows the population density for most of the island. Table 2 shows that the average population density is 226.4 people per acre. This relatively high number reflects that fact that we only include blocks with non-zero populations, and data was not available for the southern and northern most regions, where density was much lower. The densest blocks were in the Lower East Side. Other dense neighborhoods included the Upper East Side and "Hell's Kitchen" on the West Side, south of Central Park.

[10]Note in the regressions we only inlcude blocks that had non-zero populations.

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
Demographics					
Pop. Density (# per acre)	226.4	210.3	1	1281	2292
Avg. Houehold Size	4.92	2.93	0	79	2332
# German HoHs	36.7	53.8	0	434	2072
# Irish HoHs	31.3	46.6	0	401	2072
# Italian HoHs	11.0	39.1	0	492	2072
# Russian & Polish HoHs	19.4	54.6	0	466	2072
Ethnicity HHI	38.1	14.9	17.9	100	2072
Ward Density$_{1855}$	135.0	84.6	3.1	263.8	22
Sanitation District Density$_{1890}$	184.1	141.2	6.1	763.6	112
Distance, Amenity and Plating					
Distance to City Hall (miles)	3.69	2.42	0.061	8.50	2635
Distance to Broadway (miles)	0.576	0.462	0	1.87	2635
Distance to Closest River (miles)	0.414	0.258	0.008	1.01	2635
Distance to Closest El Stop (miles)	0.324	0.532	0	3.24	3180
Avg. Block Size w/in 0.25 Miles (acres)	3.50	1.01	1.43	6.93	2635
Border Park Dummy	0.143				2635
Border Park Dummy x Park Size (acres)	40.9	176.2	0	922.6	2635
Distance to Closest Park (miles)	0.235	0.206	0	1.07	2635
# of Houses of Worship with 0.5 miles	24.6	18.4	0	70	2635
Environmental					
Saltwater Area Dummy	0.223				2635
Oak Tulip Forest Area Dummy	0.692				2635
Freshwater Area Dummy	0.376				2635
Avg. Elevation (feet above sea level)	37.2	31.4	-30.6	145.4	2635
Below Sea Level Dummy	0.056				2634

Table 2: Descriptive Statistics at the block level. Sources: See Appendix A.

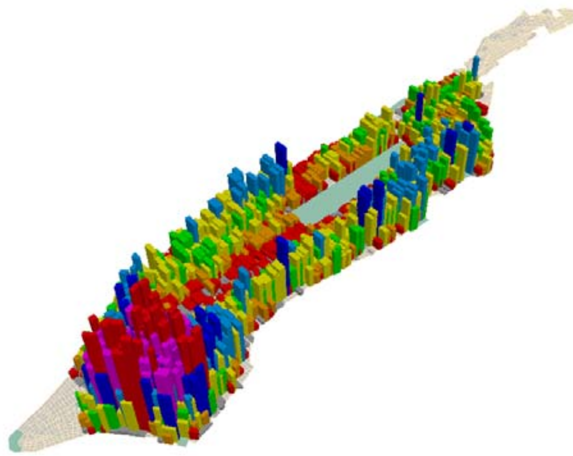

Figure 1: Manhattan population Density, 1900. Source: see Appendix A. Note data was not collected for lower wards and northern sections due to very low populations.

4.1.1 Immigration Patterns

The second half of the 19th century was a period of rapid growth for Manhattan. Between 1850 and 1900, Manhattan's population increased from 0.52 million to 1.85 million. Much of the growth was a result of European immigration. In 1900, the foreign born made up 41% of the residents. In particular, during the late-19th century, the two dominant immigrant residents were the Germans and Irish, who by 1900 made up about 43% of the foreign born population in Manhattan. The largest numbers of these two groups came in several waves between the mid-1840s to the mid-1890s. Then in the mid-1890s the Italians and Russian and Polish Jews were beginning their massive influxes, which would peak in the years before World War I. Table 3 gives the populations of the leading nationalities as of 1900.

Rank	Country	Pop.	% of F.B.	% of Pop.
1	Germany	189,720	22.3%	9.3%
2	Ireland	178,886	21.0%	8.7%
3	Russia*	130,984	15.4%	6.4%
4	Italy	103,795	12.2%	5.1%
5	Great Britain**	47,766	5.6%	2.3%
6	Hungary, Bohemia, Poland, Austria	109,458	12.9%	5.3%
	Total	760,609	89.4%	37.1%

Table 3: Foreign born population in Manhattan. *Includes Polish Russians. **English, Welch and Scotch. Source: Tables 32 & 34, *Statistics of Population*, 1900 Census, Volume 1.

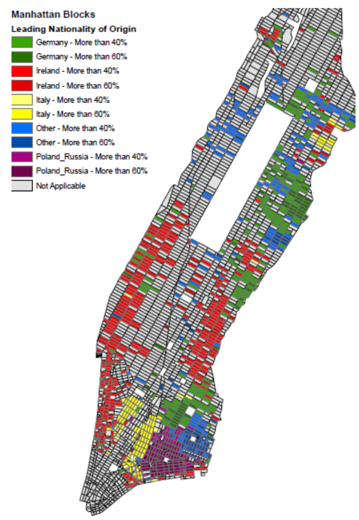

Figure 2: The distribution of heads of households based on birth country of parents. Source: See Appendix A. Note data was not collected for lower wards and northern sections due to very low populations.

Figure 2 shows the distribution of immigrant groups across the island. The map shows the relative clusters of each group on the island. The Italians were clustered in "Little Italy," the Russian and Polish Jews were concentrated in the Lower East Side. There were two concentrations of Germans. One pocket was just north of the Russian population in the Lower East Side and the second was on the Upper East Side. Finally, the Irish had two concentrations. One was on the west side from lower Manhattan to midtown and a parallel concentration on the east side.

Another perspective on immigrant locations is given in Figure 3, which shows a moving average of the density of immigrant households relative to City Hall. The graph shows distinct neighborhood concentrations of each of the four main immigrant groups. The Italians, on average, were clustered closest to City Hall; next were the Russians, who, predominantly lived about a mile away. The Irish tended to live further north; the Germans tended to live the furthest north. Also notice a small cluster of immigrants lived in the 6 to 8 mile range as well. This was the Harlem neighborhood, which had been growing as an immigrant enclave; it had been initially settled during the Dutch colonial period.

The graph illustrates that the two newest groups–the Italians and Russians–were most densely clustered, with very steep drop-offs with respect to their densities. The Irish and Germans show a more dispersed pattern. Lastly, the distance to City Hall is also likely due to the income levels of each of the groups, with the Italians and Russians having lower incomes than the Germans and Irish, who were in the country longer and had moved up on the social and economic ladder. The more northern concentration of Germans is most likely those who were in the country longer and thus had greater incomes. In summary, immigrant location appears to be based on two main criteria: desire to cluster together and income levels.

4.2 Amenities Data

4.2.1 The Natural Environment

Another set of variables relate to the environment and ecology of the island before European settlement. We aim to test how these conditions might have influenced historic settlement patterns, which, in turn, might have influenced density around the turn of the 20th century. Presumably, "favorable" environments drew early settlement, which, because of historical lock-in, was likely to draw greater population densities in the late 19th century. Conversely, unfavorable environments would likely repel populations, all else equal. The set data was collected by Eric W. Sanderson, as part of the Mannahatta Project, which recreates the ecology of the island before European settlement (Sanderson, 2009).

Before the arrival of Europeans, the island was a diverse ecosystem. Rivers, ponds, and marshes were spread throughout the land. Many species of plants and trees covered the island in forests. When Henry Hudson sailed up the river that now bears his name in 1609, some 77% of the island was covered in forest and shrubs, while another 10% were grasslands and 8% were salt and brackish marshes (Sanderson and Brown, 2007). Figure 4 gives a map of the various bodies of water on the island.

The Dutch West India Company first settled in lower Manhattan in 1624, where most of the initial residents lived together in a small village. But the Dutch also established various farms on the island, north of the settlement. Presumably, these initial farms were located where freshwater was plentiful. We would expect earlier settlement patterns to be positively correlated with the presence of freshwater. As a result, density at the turn of the 20th century was likely to be greater in these neighborhoods because of their earlier benefits, given economies of scale and the lock-in associated with economic development (Bleakley and Lin, 2012). On the other hand, areas with brackish water were likely to be avoided. These included saltwater marshes at low lying lands near the Hudson and East Rivers.[11] For each block we created one dummy variable to represent whether the land contained saltwater, and another to represent whether it contained freshwater. We expect positive coefficients for freshwater and negative for saltwater dummies.

Related to the location of water is the elevation across the island. Generally speaking, the lower part of the island is closer to sea level; going northward, the elevation tends to increase. The island was dotted with hills and ridges, that were eventually leveled or smoothed as part of the grid plan.[12] The elevation is highest on the northwestern part of the island, but drops off steeply into the low lying areas of East Harlem. During the period of rapid economic development, the low lying areas were prone to flooding, and

[11] Note that the Hudson River by Manhattan contains brackish water. Although freshwater comes from the river's head in the Adirondack Mountains, incoming tides from the ocean carry salt into the river. Also, the East River is a tidal strait, connecting New York Bay to Long Island Sound.

[12] Sanderson estimates the island had 573 hills. Allegedly, the Lenape word for the island translates into "Island of Many Hills" (Sanderson, 2009).

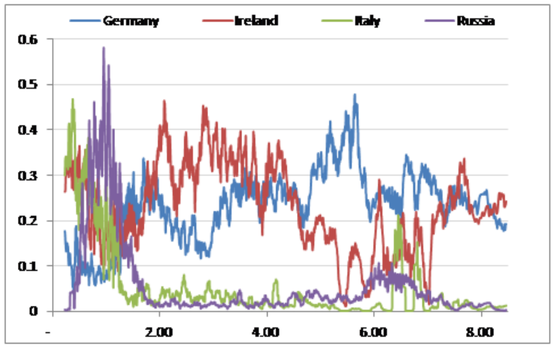

Figure 3: Fraction of each block containing a given nationality versus distance from City Hall (miles), MA(19). Source: See Appendix.

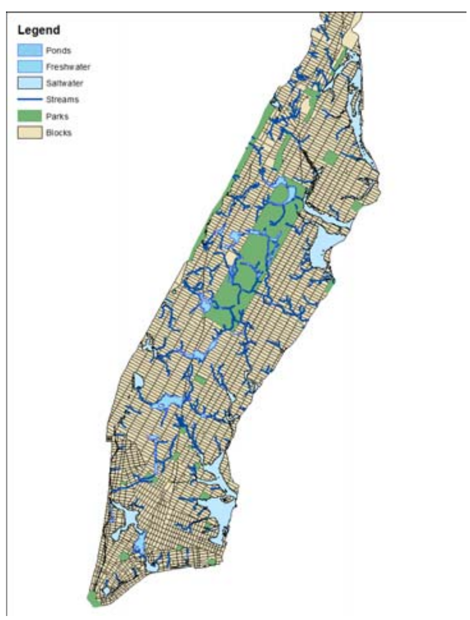

Figure 4: Wetlands, streams and ponds in pre-European settlement Manhattan. Source: See Appendix A

dampness and moisture in the cellars. These tended to have "insalubrious" effects on the neighborhoods (Tenement House Commission, 1865). We include the average elevation of the block relative to mean tide. Presumably, higher elevation blocks were more attractive because they were less likely to cause diseases. We also created a dummy variable that took on the value of one if the block was below sea level and zero otherwise. Note that blocks that were below sea level in 1609 were created out of made land. Thus this dummy variable is, in essence, a dummy variable for whether the block was created via land fill. We hypothesize a negative coefficient for blocks below sea level because they were particularly prone to disease and flooding.

The last variable is a dummy variable that took on the value one if the block was within an Oak Tulip forest, zero otherwise. These trees were important because they grew on hillsides, where the ground soil is not too sandy, and has the right amount of drainage for crops (Sanderson 2009; Sanderson personal conversation, Dec. 2012; Jorgensen, 1978). These areas of agricultural development were likely to be areas of residential and business development in the 19th century.

4.2.2 Transportation, Parks and Houses of Worship

At the turn of the 20th century, there were a few types of mass transit. On the street level were street cars running on rails embedded in the street; the cars were powered by either horse, cable, or electricity. For rapid transit, there were the Elevated Railroads (Els), which began operation in 1872 and were expanded until the 1880s. The Els contained four parallel lines, two ran up the west side of the island and two on the east side. In this paper we explore the effect of distance to the El stops on density.[13]

The next set of variables within this category relate to parks. By 1900, there were three kinds of parks. First were large parks, such as Central and Riverside Parks, which were able to accommodate many people and a variety of activities. Next, there were large squares, which were generally for strolling, but not for recreation. These included Washington Square, Union Square, and Madison Square. Finally, there were smaller neighborhood parks, of one or two square blocks, that were intended for enjoyment by residents of a particular neighborhood. Since parks were an important amenity, we look at how density was affected by the location and size of parks. The first variable we look at is the distance of each block to the closest park. Second, we created a dummy variable if the block borders a park, to test for a specific effect, since living along a park is likely to confer extra benefits (such a better view and a higher quality neighborhood). We chose to interact this variable with the size of the park, since presumably a larger park confers a greater amenity. For example, we would imagine that living adjacent to Central Park was a better amenity than living adjacent to say Tompkins Square Park, which is only a fraction of its size.

Another amenity that we investigate is the distance to houses of worship. Churches and synagogues were important for the religious, cultural and social needs of residents.

[13]Note that the New York subway did not open until 1904, so it is not directly relevant to this paper. Also note that on the east side the New York and Harlem railroad did have commuting service, but we do not explore its effects on density.

To this end, we included the number of churches or synagogues that were within a half-mile radius of each block (City of New York, 1886). The greater the number of religious institutions in a neighborhood, the greater, we hypothesize, would be the density. For this variable we make no attempt to distinguish which types of denominations were locating where. By and large, these houses of worship tended to locate in the central spine of the island. Note that the data are as of 1886. Using prior data ensures exogeneity; though the locations are from 15 years prior to our sample date, it is likely that these houses of worship remained important focal points for communities in 1900.

4.3 Locational Data

The next set of variables relates to the distance of each block to a particular location. These variables are meant to test how access to important locations affected density. First, is the distance of the block to City Hall, which was a prominent, central downtown location, and is our measure of distance to the core. Second, we look at the distance of each block to Broadway, since this street was arguably the most important shopping and employment artery, running up the spine of the island. As the Hotelling (1929) model illustrates, firms have an incentive to congregate in the "center" because of strategic interaction, and the desire to capture as large a market share as possible. Thus being close to the spine of Manhattan means being close to jobs and shopping. The question remains, however, which groups were willing to pay to be closer to Broadway?

A third variable is the distance of the block to the closest river (Hudson, East or Harlem). Because Manhattan is an island, traditionally, manufacturing, animal slaughtering, and other noxious activities were done close to the shore, because it was far away from the center and it was close to the waterways for shipping purposes. On one hand, being close to the shore can reduce commuting times for workers in shore and port-related industries. However, closeness can be a disamenity due to the negative externalities often generated by those industries.

4.4 Block Sizes

Block sizes across the island are not as uniform as one might suppose. The conventional thinking about the grid plan is that it imposed uniformity across the island. To a strong degree this is true in the sense that the street and avenue system imposed a systematic pattern, but within that, there is much variation across the island. Of the roughly 2600 blocks in our sample, the average block size is 3.4 acres (3.0 median) with a standard deviation of 1.55. The 99th percentile is 6.393 acres. The interquartile range is about 2 acres.

In our regressions we include the average size of the blocks within a quarter mile radius of each block (the block itself is excluded in the calculation), which is our measure of the degree to which the grid plan constrained builders. Our goal is to come up with a block size measure which is exogenous, hence we take the neighborhood average, and exclude the block itself.[14]

5 Empirical Results

5.1 Results without Immigrant Variables

In this section, we present the results of regressions that omit variables on immigrant status. Our interest is to first explore the effects of location, amenities, environmental, and platting variables. In Section 5.2 we include demographic and immigration variables. Table 4 presents the results of four specifications.[15] The first equation is simply a test of the standard land use model, which shows that density is negatively related to distance to the core (City Hall). On average, we see about a 7.4% drop off in density with each mile from City Hall, though the equation only has an $R^2 = 0.02$.

Equation (2) adds three other measures of location. First is the distance of each block to Broadway. Being close to the center would allow residents easier access to shopping, entertainment and employment. However, because firms are generally able to outbid residents for central locations, we would expect relatively low density on Broadway itself. If wealthy households outbid poorer households to be close to the center, we would also expect to see relatively low density near Broadway; and this is, in fact, what we see. We fit a quadratic relationship, and the results show a steady increase in density away from Broadway, but a flattening closer to the rivers.

The second locational variable is the distance of each block to the closest shore. Again we fit a quadratic relationship, finding that density peaks about one half mile from the shore. Along the shore firms outbid residents, the next blocks in, however, probably have lower density due to the presence of negative externalities; then a little further in, dock and manufacturing workers reside. The three locational variables are able to account for about 25% of the variation in population density. This suggests, on one hand, that access to key locations is important in determining density, but it also suggests that other factors are at work as well.

Equation (3) includes the amenity variables: the distance to the closest elevated railroad stop, the distance to park measures, house of worship density, and the historical ecological factors. Being close to transportation increases density as predicted. On average, density dropped by almost half with each mile away from an El stop. The park-related variables show that bordering a park reduces density, and this effect is greater for larger parks. Presumably, this is a "high utility" amenity where upper income residents outbid lower incomes ones; hence reducing density. However, being close to a park is important for all, with a negative relationship between distance and density. We also find that the more houses of worship within a half-mile radius, the greater the density. An increase in one standard deviation (18 houses) increases density by 18%.

[14] The correlation coefficient is 0.68 for the size of each block and the average within its 1/4 mile radius.

[15] We also ran spatial models with a "lagged" spatial dependent variable. Although we found evidence of spatial autocorrelation, the results were similar to the OLS models; for the sake of brevity we do not include those results.

	(1)	(2)	(3)	(4)	(5)
Distance to City Hall	−0.074	−0.123	−0.070	−0.004	0.051
	(6.71)**	(11.2)**	(3.37)**	(0.12)	(2.17)*
Distance to Broadway		1.65	2.36	2.33	1.89
		(9.58)**	(11.5)**	(11.5)**	(9.54)**
$(\text{Distance to Broadway})^2$		−0.49	−0.908	−0.873	−0.982
		(4.30)**	(7.25)**	(7.00)**	(8.08)**
Distance to Closest River		6.12	3.90	3.94	3.17
		(15.6)**	(8.27)**	(8.34)**	(6.81)**
$(\text{Distance to Closest River})^2$		−6.06	−4.37	−4.35	−3.52
		(14.9)**	(10.0)**	(9.96)**	(8.27)**
Distance to Closest El Stop			−0.556	−0.675	−0.463
			(2.80)**	(3.34)**	(2.31)*
Border Park Dummy x Park Size			−0.0005	−0.0005	−0.0006
			(4.34)**	(4.09)**	(5.57)**
Distance to Closest Park			−0.884	−0.855	−0.884
			(5.25)**	(5.01)**	(5.4)**
# Houses of Worship w/in 0.5 Miles			0.010	0.008	0.002
			(4.22)**	(3.44)**	(0.79)
Saltwater Area Dummy			−0.013	−0.018	0.058
			(0.20)	(0.26)	(0.88)
Oak Tulip Area Dummy			0.191	0.196	0.176
			(3.28)**	(3.35)**	(3.10)**
Freshwater Area Dummy			0.083	0.087	0.079
			(1.81)	(1.89)	(1.78)
Elevation			−0.0002	0.0001	−0.000
			(0.22)	(0.08)	(0.03)
Below Sea Level Dummy			−0.595	−0.617	−0.560
			(3.61)**	(3.68)**	(3.4)**
Avg. Block Size w/in 0.25 Miles			0.090	0.082	0.027
			(2.99)**	(2.68)**	(0.9)
$\text{Ln}(\text{Ward Desnity}_{1855})$				0.113	
				(2.61)**	
$\text{Ln}(\text{SD Density}_{1890})$					0.519
					(11.4)**
Constant	5.2	3.5	3.3	2.7	1.2
	(100.5)**	(37.5)**	(20.2)**	(10.0)**	(4.8)**
Observations	2292	2292	2291	2291	2291
R^2	0.02	0.25	0.29	0.29	0.34
$\bar{R}^2$	0.02	0.24	0.29	0.29	0.33

Table 4: Block-level regressions. The dependent variable is ln(Pop.Density). Absolute value of robust t-statistics below coefficient estimates. *Statistically significant at the 95% confidence level; **Statistically significant at the 99% confidence level.

Generally speaking, the environmental variables show the predicted signs, though statistical significance varies. Oak Tulip areas and freshwater areas are positively associated with density, suggesting that these areas were attractive to early European and American residents – and this attractiveness persisted until the 19th century. The saltwater regions are associated with negative density in two out of three specification (though not statistically significant in all specifications). These regions tend to be by the banks of the rivers, where port activity most likely dominated in the 19th century.[16] The two elevation-related variables have negative signs, but the elevation variable is not statistically significant – suggesting, elevation, in an of itself, was not an important determinant. But we discuss this variable more in the next section. The blocks below sea level show a large and statistically significant negative effect on density, as was expected, given how unhealthy low lying areas were; and further they tended to be on made land, near the rivers.

Equation (3) also includes the average block size, finding a positive relationship, which is the opposite of our expectation. However, we leave a more detailed discussion of this variable for the next section, after we include demographic variables–since this estimate may suffer from an omitted variable bias, since the smallest blocks are located in lower Manhattan, where there are the greatest immigrant concentrations. Overall, the inclusion of these additional variables only increases the R^2 by about four percentage points to 0.29, which suggests a modest role for neighborhood amenities.

Finally, equations (4) and (5) provide historical neighborhood density measures. Equation (4) includes the log of the ward density in 1855; equation (5) includes the log of the density at the sanitation district level in 1890 (Billings, 1894). Each ward was subdivided into one or more sanitation districts (descriptive statistics are in Appendix B). Roughly speaking each SD can be considered a large neighborhood. These variables can be thought a type of lag dependent variables. The goal is to see how earlier density patterns affect density in 1900 and how the inclusion of these variables affect the other coefficient estimates. First we see that earlier settlement patterns are positively related to 1900 density, as would be expected. Immigrants would be attracted to places were the housing stock is older and hence less expensive (and also contained the institutions and businesses that were attractive to these groups). Second, the inclusion of the ward density in 1855 does not generally change the other coefficients, except for the distance to City Hall variable. Now we see that, controlling for prior density, the effect seems to be zero. This suggests that the distance variable is a proxy for earlier settlement patterns. Equation (5) includes the log of density at the Sanitation District (SD) level in 1890. By and large, most of the variables retain their signs and statistical significance.

The inclusion of lagged density for 1890 changes the sign of the distance to City Hall coefficient, which is now positive. This finding suggests the island in the 1890s was, in some sense, decentralizing – most notably with the rise of dense neighborhoods around and north of Central Park. Therefore, controlling for historical density patterns, the increase

[16] Note that the saltwater dummy variable remains statistically insignificant when the distance to the shore variable is removed. This suggests that the locations of former marshlands were not a strong factor in the location of residential dwellings.

in density was occurring further away, most likely due to rising incomes and increased use of rapid transit.

5.2 Results with Immigration Variables

This section presents results with the inclusion of immigration and demographic variables. In particular, we include the average household size for each block (total population divided by number of households), the number of household heads (HoHs) with parents from different countries, and a measure of immigrant diversity as the block level.[17] Our diversity measure is the HHI equal to the sum of the HoH immigrant shares squared for each block.

Specifically, we created five nationality categories for the counts of the household heads: Italian, Irish, Russian-Polish, German and one category for the rest.[18] A note is order about specification. On the right hand side we included the log of the *number* of heads of households in one of the five categories. We chose this variable rather than the percentage of each group because we are interested in how immigration itself influenced the spatial distribution of density. The percentage, on the other hand, is a measure of relative composition, i.e., a measure homo- or heterogeneity. We return to composition effects when we investigate the Lower East Side in more detail, in Section 5.4.

Because the variable was the count of head of households, rather then total number of people from a region or nation, it is less likely to have a direct, mechanical influence on the overall density measure. Having said that, because having an HoH variable is likely to be endogenous with respect to the dependent variable, we have employed two-stage and three-stage least variables, using lagged immigration data, to create an exogenous measure of the number of immigrants on the block.

Table 5 presents the results. Equation (1) is estimated via ordinary least squares and includes the distance-related variables, and the demographic variables (but no amenity, environmental or block size variables). The distance variables retain their signs and significances. The demographic variables show the expected signs. The omitted group for the immigrants are those heads of households with native parents. We find that density is higher for Italians, Russian/Polish and the combined group, compared to German and Irish dominated blocks, as would be expected. We find a positive relationship between family size and density, as expected.[19] The inclusion of demographics dramatically increases the R^2, as would also be expected.

[17] Specifically, the variable is the birth country of the parents of the head of household. This was the information collected by the census and amalgamated at the block level for the Tenement House Report (1903).

[18] We assume that average household size is exogenous. Further, we put Russian and Polish immigrants together in one category since most likely they shared a common Jewish heritage, and these groups were lumped as one in 1890 Census Bureau's "Vital Statistics" report (Billings, 1894).

[19] We limit the data in the regressions to those observations with average family size less than 20 individuals. Note that the 99th percentile for this variable is 10.0, thus we exclude only a few of the blocks with extreme outliers. It's not clear from the tables if these outliers are errors, or perhaps they are due to the presence of lodgers or extended families.

Variable	(1) OLS	(2) IV	(3) 3SLS	(4) OLS	(5) 3SLS
Distance to City Hall	−0.046	−0.020	−0.016	−0.005	0.009
	(6.39)**	(2.10)*	(1.96)*	(0.44)	(0.79)
Distance to Broadway	0.318	0.415	0.342	0.380	0.311
	(3.64)**	(3.89)**	(3.73)**	(3.82)**	(2.81)**
(Distance to Broadway)^2	−0.187	−0.209	−0.187	−0.257	−0.212
	(3.34)**	(3.14)**	(3.37)**	(4.40)**	(3.50)**
Distance to Closest River	1.45	1.93	1.38	0.664	0.714
	(7.11)**	(7.36)**	(6.61)**	(3.02)**	(3.30)**
(Distance to Closest River)^2	−1.35	−1.98	−1.31	−0.609	−0.633
	(6.68)**	(7.40)**	(6.05)**	(2.95)**	(2.97)**
Distance to Closest El Stop				−0.203	−0.204
				(2.11)*	(2.05)*
Border Park Dummy x Park Size				−0.0001	−0.0001
				(1.64)	(1.88)
Distance to Closest Park				−0.301	−0.245
				(3.88)**	(3.33)**
# Houses of Worship w/in 0.5 Miles				0.002	0.002
				(2.16)*	(1.63)*
Saltwater Area Dummy				−0.096	−0.081
				(3.18)**	(2.72)**
Oak Tulip Area Dummy				−0.072	−0.069
				(2.72)**	(2.72)**
Freshwater Area Dummy				0.016	0.011
				(1.05)	(0.73)
Elevation				0.001	0.001
				(2.24)*	(1.82)
Below Sea Level Dummy				−0.215	−0.241
				(3.06)**	(3.91)**
Avg. Block Size w/in 0.25 Miles				−0.203	−0.176
				(11.78)**	(7.88)**
ln(Avg. Household Size)	0.893	0.513	0.887	0.961	0.980
	(8.35)**	(4.62)**	(12.4)**	(9.17)**	(13.04)**
ln(German HoHs)	0.085	−0.032	0.003	0.138	0.050
	(5.59)**	(1.04)	(0.09)	(9.21)**	(1.70)
ln(Irish HoHs)	0.108	0.058	0.075	0.143	0.157
	(9.95**	(2.64)**	(3.19)**	(13.56)**	(5.95)**
ln(Italian HoHs)	0.131	0.174	0.159	0.122	0.124
	(13.6)**	(7.43)**	(7.22)**	(13.56)**	(6.03)**
ln(Russian/Polish HoHs)	0.147	0.295	0.254	0.105	0.171
	(14.6)**	(13.8)**	(13.4)**	(10.72)**	(8.59)**
ln(Other Nat. HoHs)	0.312	0.157	0.201	0.326	0.302
	(17.10)**	(4.37)**	(5.60)**	(21.2)**	(9.10)**
HHI	0.011	0.001	0.009	0.012	0.012
	(8.33)**	(0.39)	(6.52)**	(10.9)**	(7.40)**
Constant	2.14	2.81	1.83	1.86	1.88
	(11.1)**	(10.2)**	(8.66)**	(8.48)**	(8.39)**
Observations	2053	2053	2053	2053	2053
R^2	0.76	0.69	0.71	0.80	0.79
$\bar{R}^2$	0.76			0.80	

Table 5: Block-level regressions. The dependent variable is ln(Pop. Density). Absolute value of robust t-statistics below coefficient estimates in equation (1), (2) and (4). z-statistics below the estimates of the others. *Statistically significant at the 95% confidence level; **Statistically significant at the 99% confidence level.

Equation (2) contains the same variables but uses instrumental variables for each immigrant group. For instruments, we used 1890s variables. In particular, the first stage regressed the immigrant variables on the log of the number of German, Irish, Italians, and Russian/Polish immigrants, and Blacks in each Sanitation District as of 1890. Additional first stage controls included the percent of the population in each SD that was less than five years old, the SD death rate, and the ward density in 1855 (descriptive statistics are in Appendix B). All of these instruments are meant to predict immigrant density patterns based on prior decisions of immigrants, and some features of the neighborhoods. Neighborhoods with relatively high number of children might be considered a kind of neighborhood amenity, i.e., these neighborhoods were particularly "family friendly." Neighborhoods with high death rates might suggest the opposite. Ward density in 1855 is meant to capture the relative age of the housing stock in general.

The robust χ^2-statistic for the endogeneity tests had a p-value of 0.00, suggesting endogeneity is an issue. A note is in order about the overidentification test for the validity of the instruments. When we included all of the instruments except the death rate and the percent of the population less than five years old, the p-value for the overidentification test was $p = 0.58$,but when included them the p-value was $p = 0.002$. In the end we decided to include the two extra instruments for two reasons. First, since they relate to data a decade prior they are *a priori* exogenous. Second, they are interesting variables to investigate when using a three-stage least squares regression (results discussed below).[20] Also note that the F-statistic for all first-stage equations greater than 47.0, indicating strong instruments.

Equation (3) is the same equation but is estimated with three stage least squares, using the same first-stage variables as above; again, we don't see large differences in the coefficient estimates. Equation (4) is the "full" equation, estimated by OLS. That is, equation (4) includes all of the location, amenity, block size and environmental variables, in addition to the demographic variables. With the full set of controls, we see that the saltwater dummy variable is now negative, as we would expect, but the Oak Tulip variable switches signs to negative. This suggests that controlling for demographic patterns, these regions no longer retain their attractiveness. The freshwater variable remains positive, as does the elevation variable. The below-sea-level dummy remains negative. Furthermore, block size now has a negative effect as well. This would suggest that the small blocks placed constraints on builders, who, in turn, increased the density of their structures. Equation (5) is the same equation but estimated using three stage least squares. Again, we see very similar results as those with OLS. The only major difference is the reduced size of the distance to City Hall variable. Finally, assuming the HHI is exogenous, it shows that increases in concentration (greater HHI values) are also associated with greater density.

Interestingly, across the city, the most concentrated blocks were those immigrants from other nationalities rather than the "big four." In some specifications, such as in equation (5), the German coefficient is negative, suggesting less density than natives. This coefficient, however, is the least stable across specifications, compared to other nationalities.

[20] Note for the sake of brevity we do not present results from the first-stage regressions. They are avaiable upon request.

In general, among the big four, the Russians/Polish and Italian groups seem to have the largest densities, as would be expected.

5.3 Immigrant Preferences

The results of Table 5 suggest the question, what drove immigrants to choose some locations over others? We know that time in the country and income was an important factor for how far they lived from downtown. We also know that, by and large, those regions in between Broadway and the rivers were also high concentrations of density, since they were in a kind of "sweet spot"–not too far away from important locations (Broadway and the rivers) yet they remained affordable for these classes. Further, we know that density was affected by amenities, such as the location of parks, the ELs and the least desirable marshlands.

But this section presents the results from the three stage least squares regressions to offer some evidence on the preferences of the immigrants themselves, and what demographic characteristics were desirable to them in general. Above, Table 5 equation (5) showed the result of only the density equation from the three-stage least squares regression, but here we show the results of the other equations in the system. In particular, the dependent variables are the log of one plus the number of heads of households for Italians, Germans, Russian & Polish, Irish and the other groups, respectively. Based on the results, we can draw some conclusions about the determinants of immigrant locations at the turn of the 20th century.

Table 6 presents the results. Each equation investigates how immigrant location decisions in 1900 were determined by immigrant (and Black) location decisions in 1890, and two neighborhood demographic variables–the percent of sanitation district residents less than five years old, and the death rate (per 1000 people) of the SD in 1890. By including the number of very young children, we aim to test for how different groups might have perceived these neighborhoods as good for families. Death rates can affect the location decision since, presumably, it is also a (negative) measure of the quality of life. Finally, we control for the age of the housing stock and prior density by including the ward density as of 1855.

The results show how each of the five groups were attracted, or not, to different neighborhoods based on the composition of the neighborhood in the past. For all groups we see that each group was attracted to neighborhoods with a high prior concentration of their own group, as would be expected. The coefficients for the Italian and Russian/Polish groups are the lowest, however. This probably reflects the fact that their massive influxes into the city began in the 1890s.

Looking at the signs of the right hand side coefficients can suggest which immigrant groups were "attracted" or "repelled" by the other immigrant groups. The Germans seem to have preferred living apart from the other groups, as most of the coefficients for the other groups are negative. However, the Germans seemed to be attractive to other groups, as the right hand side coefficients for them were all positive. The Irish seem to have been "repellents" as all of their coefficients are negative. We see similar effects for the Italians and Russians. We see a more mixed picture for Blacks, with some negative and

Variable	German	Irish	Italian	Russ./Pol.	Other
ln(# Germans in SD_{1890})	0.998	0.154	0.114	0.278	0.258
	(19.3)**	(2.66)**	(2.22)*	(5.87)**	(5.09)**
ln(# Irish in SD_{1890})	−0.050	1.10	−0.138	−0.230	−0.248
	(1.11)	(22.1)**	(3.13)**	(5.66)**	(5.69)**
ln(# Italians in SD_{1890})	0.009	−0.137	0.392	−0.078	−0.016
	(0.33)	(4.59)**	(14.9)**	(3.22)**	(0.61)
ln(# Russian/Polish in SD_{1890})	−0.229	−0.202	−0.093	0.511	−0.106
	(10.5)**	(8.25)**	(4.34)	(25.69)**	(4.98)**
ln(# Other in SD_{1890})	−0.032	0.038	−0.154	−0.125	0.752
	(0.46)	(0.49)	(2.25)*	(1.96)	(11.1)**
ln(# Black in SD_{1890})	−0.065	0.036	0.055	−0.067	−0.089
	(2.57)*	(1.25)	(2.14)*	(2.84)**	(3.57)**
% Pop. < 5 y/o in SD_{1890}	0.095	0.123	0.070	0.110	0.145
	(7.84)**	(9.00)**	(5.83)**	(9.88)**	(12.25)**
Death Rate in SD_{1890}	−0.026	−0.009	−0.006	−0.039	−0.037
	(4.63)**	(1.35)	(1.11)	(7.47)**	(6.71)**
ln(Ward $Density_{1855}$)	−0.01	0.217	0.242	0.084	0.041
	(0.37)	(10.0)**	(12.7)**	(4.75)**	(2.15)*
Constant	−3.33	−8.10	−0.573	0.081	−2.51
	(8.13)**	(17.6)**	(1.41)	(0.22)	(6.27)*
psuedo -R^2	0.37	0.29	0.32	0.57	0.34

Table 6: Results from three-stage least squares. Dependent variables: ln(1+HoH Count) for 1900. Absolute value of z-scores below estiamtes. Sample size is 2053 for all equations. *Statistically significant at the 95% confidence level; **Statistically significant at the 99% confidence level.

some positive coefficients. In summary, the results suggest that while the Germans were not interested in living side-by-side with the other immigrants, living near Germans was considered a neighborhood amenity of sorts. In other words, the interaction of immigrant groups appears to have been an important determinant of spatial structure.

All immigrant groups seemed to have preferred neighborhoods with many young children, and they, by and large avoided neighborhoods with high death rates. The evidence is also consistent with the fact that immigrant groups had a certain amount of locational choices; and preferred neighborhoods with better bundles of amenities.

Finally, the coefficient signs for the 1855 ward density provide evidence on the relative income of the groups. Except for the Germans, 1855 ward density was positively related density in 1900 for the various immigrant groups. Interestingly, the coefficient was the largest for the Irish and Italian groups. By and large, the result suggest that the preferences of the different groups had an important effect on urban spatial structure. Each group generally had different income levels, chose different occupations, and had different cultures. The way these groups interacted, ultimately effected the different neighborhood densities and the locations of the other groups. These results, however, suggest there is a greater need for further investigations into how immigrant and cultural dynamics can influence the location of economic activity across the city.

5.4 The Lower East Side

The conventional understanding of the Lower East Side is that the combined forces of massive immigration, its historical place as a tenement district, and its close location to jobs in lower Manhattan and Broadway created unprecedented urban density. By and large, our regression results show this to be true. However, even within the LES there was a fair degree of variation of extreme density from one block to the next. Figure 5 shows the location of the highest density blocks within the LES. In the section we aim to understand the factors drove the variation across this particular region of the city, where we define the Lower East Side as comprising wards 7, 10, 11, 13 and 17. To investigate this we include several new variables. More specific information about these variables is contained in the appendices.

First we aim to see how the physical stock of the blocks influenced density. We created one variable by counting the number of non-housing structures on each block, as of 1898. Non-housing structures generally included factories or horse stables.[21] Second, as a control for the quality of the housing stock, we counted the number of rear tenement apartments on each block. These buildings were considered a promoter of disease and despair among the immigrant communities. Jacob Riis (1890) made rear tenements famous in his photographs which presented images of deep privation. We hypothesize that the more rear tenements the greater the population, since they were likely to attract the

[21] Practically all buildings in the Lower East Side with street frontage, either tenements or non-tenements, had some non-housing use. Most tenements, for example, had stores on the ground floors. The data that we used for non-housing structures are the ones that were specifically designated as having no residential use at all.

Figure 5: Location of extremely dense blocks within the Lower East Side. Shaded blocks have density of at least 900 people per acre. Colored lines are ward boundaries. Source: See Appendix A.

poorest of the poor. These variables were counted from a Manhattan atlas from 1898 (Bromley and Bromley, 1899).

As a measure of a neighborhood amenity, we created a variable that gave the distance of each block to the closest synagogue as of 1884 (Edwards, 1884), since the neighborhood had a large concentration of Jews. Next we included a dummy variable that takes on the value of one if the block has an irregular shape (i.e., is not perfectly square or rectangular); zero otherwise. We hypothesize that irregular blocks would have less density since it would be more difficult or expensive to build housing on these blocks. Lastly we control for compositional effects (the percentage of immigrant groups living on each block) to see if relative concentrations were causal.

Table 7 presents the regression results, where the dependent variable is the log of population density on each block. Equation (1) is estimated via OLS; Equation (2) is estimated via three stage least squares, where the instrumental variables is from the 1890 sanitation districts: the percentage of each of the immigrant groups and blacks, the percentage of population less than five year old, the death rate, and the log of 1855 ward density (descriptive statistics in Appendix B).

Looking at the new variables, we see that the evidence of the shape of the block is inconclusive, as it negative and significant in the OLS regression, but appears to be zero in the three-stage least squares regression. The number of non-housing structures and rear tenements do appear to be quite important. The coefficient for non-housing structures is negative, with an elasticity of about 0.15. The number of rear tenements was also important, showing the effect to be positive, with an elasticity of about 0.16-0.17. It appears that distance to a synagogue was important, with a negative effect as the distance increases, though the magnitude and significance vary across regressions. Finally looking at the percentage of each groups suggests that the LES's density was determined more by general immigration, patterns rather than the composition of any one group on

	(1)	(2)
Variable	OLS	3SLS
Distance to Broadway	1.08	−0.29
	(2.52)*	(0.34)
Distance to City Hall	−0.125	0.225
	(0.34)	(0.29)
Distance to Closest River	5.79	4.53
	(3.11)**	(2.04)*
(Distance to Closest River)^2	−4.57	−4.60
	(2.89)**	(2.54)*
Irregular Block Dummy	−0.135	0.02
	(1.44)	(0.18)
ln(# NonHousing Structures)	−0.15	−0.15
	(3.39)**	(2.57)*
ln(# Rear Tenements)	0.18	0.16
	(5.25)**	(2.71)**
Distance to Closest Synagogue	−1.04	−0.39
	(2.75)**	(0.69)
ln(Avg. Houshold Size)	−0.59	−0.905
	(1.48)	(1.36)
Below Sea Level Dummy	−0.60	−0.79
	(1.19)	(1.83)
ln(SD Density$_{1890}$)	−0.40	−0.51
	(2.26)*	(2.75)**
% Russian/Polish	0.030	0.130
	(5.02)**	(3.80)**
% Italian	0.031	0.12
	(4.56)**	(3.79)**
% Irish	0.018	0.143
	(2.44)*	(4.11)**
% German	0.017	0.116
	(2.32)*	(3.98)**
% Other	0.023	0.144
	(2.68)**	(3.91)**
Constant	5.30	−2.73
	(4.01)**	(1.10)
R-squared	0.63	
Ward Dummies P-value	0.001	0.005

Table 7: Regression results for ln(Pop. Density) for each block in the Lower East Side (wards 7, 10, 11, 13 and 17). Number of observation is 295. See Appendix for data sources. Absolute value of robust t or z stats. below estimates. **Stat. sig. at 1% level; * Stat. sig. at 5% level.

a block.

Interestingly, a few of the variables from the original regressions show a change of sign when we focus only on the Lower East Side. In particular the 1890 SD density, and household size are now negatively related to density. We leave investigating these results for future work; but they suggest a more set of more complex intra-neighborhood dynamics.

6 Conclusion

This paper has explored the causes for population density variation on Manhattan Island in the year 1900. Manhattan presents a unique case study, given that it was a draw for millions of immigrants, and there was a high degree of density variation across the island. In the Lower East Side, some blocks contained over 1200 people per acre, while the overall density of the city was closer to 150 people per acre. The literature on urban spatial structure tends to focus on the effects of distance and income on density across metropolitan regions. Less work has investigated the patterns of density within cities. The historical decisions made by households and builders continues to be relevant today, given that current land use patterns are generally similar to those of a century ago, with distinct tenement, upper income apartment districts, and business and commercial areas.

We investigate the effects of several types of variables on density. We aim to understand how access to important locations, distance to amenities, early environmental factors and immigrant patterns affect population density across the island. We find that while the locational variables are important, they only explain about a quarter of the variation. Investigating the role of natural and produced amenities shows that, while they can explain some of the variation, their effects are small relative to immigrant residential patterns. We also find that in addition to time in the country and desire to cluster, the preferences of immigrant groups to live near or far away from other groups also affected density across the island. In particular the Germans appeared to have been a positive "amenity" for the other groups. Further we find evidence for the role of the grid plan, with larger blocks being less dense, all else equal.

Lastly we also find that in the Lower East Side, its extreme density across blocks can also be explained by the building types that impacted the available housing. We find that the number of rear tenements was positively related to density, while the number of horse stables and non-residential buildings negatively determined density. Since immigrants preferred living among themselves, if one block had a particularly low amount of housing it would likely increase density in near by blocks.

In this paper, we do not have data on housing prices. Future work can test how housing prices influence density and land use patterns. Other work can compare urban density patterns across cities to see which factors are universal and which are distinctly local. Finally, future work can compare density on each block today to density from 1900 to see if there still remains a strong relationship, even after years of urban depopulation and decentralization.

A Data Sources and Preparation

Population Density: First Report of the Tenement House Department of the City of New York, Volume II (1903). The population, density and land area of each block is given in the maps included in the volume. Note that the Report contained its own block numbering system, which is different than the more commonly used tax block ID numbers. As part of the data processing, we also matched each block to the tax block ID number.

Foreign-born Status and *Parentage of Head of Household*: *First Report of the Tenement House Department of the City of New York,* Volume II (1903), Tables on pages 31-103. Note for Figures 2 and 3, we took the percentage of each group. For Figure 2, blocks with less than 40% for any all of the foreign groups were denoted "NA."

1900 *Foreign Born Population*: 1900 Census, *Statistics of Population,* Volume 1. Tables 32 & 34.

Map of Manhattan: Manhattan map here is a digitized map of Manhattan circa 1900. The maps created by first downloading the Manhattan blocks shapefile from the Bytes of the Big Apple website, hosted by the NYC Department of City Planning (in 2010). Then digitized maps of Manhattan Land Books from 1911 and 1898, (posted on the New York City Public Library website) were imported into ArgGIS and aligned with the current blocks. When the 2010 blocks were different than the earlier blocks, the shapes were redrawn to conform with the blocks circa 1900.

Tenement Heights: Veiller (1903b).

Manhattan Waters: The geographic location of the various water types and ecology (including Ponds, Streams, Oak Tulip Forests, Freshwater, and Saltwater) were taken from data collected by the Wildlife Conservation Society's Mannahatta Project. The locations of ponds were taken from a file named "mannahatta_ponds.zip." The locations of streams were taken from the file "mannahatta_streams.zip". The locations of Oak Tulip forests were taken from a file named "ecocomm.zip." The locations of freshwater were aggregated from a file named "ecocomm.zip" to include all of the following ecological types: "Coastal plain pond community," "Eutrophic pond community," "Deep emergent marsh," "Shallow emergent marsh," "Shrub swamp," and "Red maple-hardwood swamp." The locations of saltwater were aggregated from a file named "ecocomm.zip" to include all of the following ecological types: "Marine eelgrass meadow", "Marine intertidal mudflat community," "Marine gravel/sand beach community", "Low salt marsh," "High salt marsh," "Coastal salt pond community," and "Brackish tidal marsh." The files are available at: http://welikia.org/download/scientific-data/

Average Elevation: : The elevation, in feet above mean tide at Sandy Hook, was averaged for each block from the Digital Elevation Model prepared by Professor Eric Sanderson. The file was made available by Sanderson on request.

Average Household Size: First Report of the Tenement House Department of the City of New York Volume II (1903) gives the population of each block, as well as the number of households. The latter was divided by the former to calculate the average.

Distance to City Hall: The shortest straight line distance from each block to City Hall, measured in ArcGIS.

Distance to Broadway: The shortest straight line distance from each block to the nearest point on Broadway, measured in ArcGIS.

Distance to Nearest Park: The shortest straight line distance from each block to the nearest park, measured in ArcGIS.

Borders a Park: If ArcGIS reported zero distance between a block and a park, then the variable was given 1, 0 otherwise. Zero distance occurs when two blocks or a block and park share a border, which is set in the middle of each street.

Distance to Nearest Shore: The shortest straight line distance from each block to the nearest point on either the Hudson or East Rivers, measured in ArcGIS.

Block Size: The size of most blocks is given in the maps in First Report of the Tenement House Department of the City of New York Volume II (1903). Of the 3,181 blocks studied, 71% were included in the Report. For the remaining 29%, we first regressed block size from the Tenement House report on block size calculated in ArcGIS and used predicted values to calculate block sizes from the ArcGIS data.

Elevated Railroads: "Map of the Elevated Railroads of New York City" retrieved from: http://www.loc.gov/resource/g3804n.rr004540/. The distance from each block to the stops on the elevated railroad were measured in ArcGIS.

Sanitation District Density: From *Vital Statistics of New York City and Brooklyn* (Billings, 1890). The density is persons per acre excluding parks and cemeteries from Table 76, on page 76.

Sanitation District Demographics: Vital Statistics (1894), Table 78.

Sanitation District Birthplace of Mothers: *Vital Statistics (1894)*, Table 79.

Sanitation District Death Rate: Vital Statistics (1894), Table 84.

Number of non-housing Units, Number of Rear Tenements and Shape of Block. Bromley and Bromley (1899).

Synagogues on the Lower East Side. Edwards (1884).

B Additional Descriptive Statistics

Here we present descriptive statistics for the additional variables used in sections 5.3 and 5.4.

Variable	Mean	St. Dev.	Min.	Max.	Nobs.
		Sanitation District Data			
SD # German	3385.6	3332.4	40.0	16500.0	112
SD # Irish	3418.9	2367.3	117.0	11271.0	112
SD # Italian	473.0	1241.1	0.0	8280.0	112
SD # Russian/Polish	713.8	2620.6	1.0	21311.0	112
SD # Other	2187.3	1765.4	258.0	12546.0	105
SD # Black	227.2	391.9	2.0	2231.0	112
SD % German	23.0	13.5	5.4	76.1	112
SD % Irish	28.9	12.8	1.0	60.7	112
SD % Italian	4.1	9.7	0.0	49.5	112
SD % Russian/Polish	3.4	9.7	0.0	69.9	112
SD % Other	16.4	8.0	5.3	51.3	105
SD % Black	2.0	2.9	0.0	18.2	112
SD % Pop. $<$ 5 y/o	10.1	3.4	1.1	16.4	105
SD Death Rate (per 1000)	29.6	8.7	9.2	59.2	112
		Lower East Side Data			
# Synagogues within Half Mile	41.2	12.8	12	70	477
Irregular Dummy	0.2	0.4	0	1	298
# Nonhousing Structures	2.7	2.6	0	14	298
# Rear Structures	4.1	4.7	0	32	298

Table 8: Additional Descriptive Statistics. Sources: See Appendix A.

References

[1] Anas, A., Arnott, R. and Small, K. A. (1998). "Urban Spatial Structure." Journal of Economic Literature, 36, 1426-1464.

[2] Atack, J., and Margo, R. A. (1998) "Location, location, location! The Price Gradient for Vacant Urban Land: New York, 1835 to 1900." The Journal of Real Estate Finance and Economics 16(2), 151-172.

[3] Barr, J. Tassier, T. and Trendafilov, R. (2010). "Bedrock Depth and the Formation of the Manhattan Skyline, 1890-1915." Journal of Economic History, 71(4), 1060-1077.

[4] Billings, J. S. (1894). Vital Statistics of New York and Brooklyn, Covering a Period of Six Years Ending May 31, 1890. Washington DC: U.S. Census Bureau.

[5] Bleakley, H. and Lin, J. (2012). "Portage and Path Dependence." Quarterly Journal of Economics, 127, 587-644.

[6] Bromley, G. W. and Bromley, W. S. (1899). Atlas of the City of New York, Borough of Manhattan. Philadelphia: G. W. Bromley.

[7] Brueckner, J. K. (1987). "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model. Handbook of Regional and Urban Economics, Vol. II. Ed. Edwin S. Mills.

[8] Brueckner, J. K., Thisse, J-F, Zenou, Y. (1999). "Why is Central Paris Rich and Downtown Detroit Poor? An Amenity-based Theory." European Economic Review, 43, 91-107.

[9] Burchfield, M., Overman, H. G., Puga, D. and Turner, M. A. (2006). "Causes of Sprawl: A Portrait from Space." The Quarterly Journal of Economics, 121(2),587-633.

[10] Citizens' Association of New York (1866/1970). Report of the Council of Hygiene and Public Health. Reprint Edition. New York: Arno Press.

[11] (The) City of New York: A Complete Guide (1886). New York: Taintor Brothers, Merrill & Co.

[12] Cutler, D. M. and Glaeser, E. L (1997). "Are Ghettos Good or Bad?" The Quarterly Journal of Economics, 112 (3): 827-872.

[13] Davis, D. R., and Weinstein, D. E. (2002). "Bones, Bombs, and Break Points: The Geography of Economic Activity." American Economic Review, 92(5), 1269-1289.

[14] DeForest, R. W. and Veiller, L. (eds.) (1903). The Tenement House Problem. New York: The MacMillen Company, Volumes 1 and 2.

[15] Edwards, R. (1884). New York's Great Industries. New York. Historical Publishing Company.

[16] Flagg, E. (1884). "The New York Tenement-House Evil and Its Cure." Scribner's Magazine, 16, 108-117.

[17] Fujita and Ogawa (1982). "Multiple Equilibria and Structural Transition of Non-monocentric Urban Configurations." Regional Science and Urban Economics, 12(2), 161-196.

[18] Glaeser, E. L., Kahn, M. and Rappaport, J (2008). "Why do the Poor Live in Cities? The Role of Public Transportation." Journal of Urban Economics. 63(1): 1-24..

[19] Glaeser, E. L. and Khan, M. E. (2004). "Sprawl and Urban Growth." Handbook of Regional and Urban Economics, Vol. 4, Chapter 56, Eds. J. Vernon Henderson and Jacques-François Thisse.

[20] Hatton, T. J. and Williamson, J. G. (1994). "What Drove the Mass Migrations from Europe in the Late Nineteenth Century?" Population and Development Review, 20(3), 533-559.

[21] Helsley, R. W. and Sullivan, A. M. (1991). "Urban Subcenter Formation." Regional Science and Urban Economics, 21, 255-275.

[22] Hotelling, H. (1929). "Stability in Competition." The Economic Journal, 39(153), 51-57.

[23] Jorgensen, N. (1978). A Sierra Club Naturalist's Guide to Southern New England. San Francisco: Sierra Club Books.

[24] Margo, R. A. (1996). "The Rental Price of Housing in New York City, 1830-1860." The Journal of Economic History 56(3), 605-625.

[25] McDonald, J. F. (1989). "Econometric Studies of Urban Population Density: A Survey." Journal of Urban Economics, 26, 361-385.

[26] Muth, R. F. (1969). Cities and Housing: The Spatial Pattern of Urban Residential Land Use. Chicago: The University of Chicago Press.

[27] Pratt, E.E. (1911). Industrial Causes of Congestion of Population in New York City. New York: Columbia University Press.

[28] Riis, J. (1890). How the Other Half Lives. New York: Charles Scribner's Sons.

[29] Saiz, A. (2010). "The Geographic Determinants of Housing Supply." Quarterly Journal of Economics, 125(3) 1253-1294.

[30] Report of the Tenement House Committee (1895). Albany.

[31] Report on Vital and Social Statistics in the United States at the Eleventh Census, 1890 (1896):. Washington DC: U.S. Census Bureau.

[32] Tenement House Department (1903). First Report, Volumes I and II. New York.

[33] Sanderson, E. W. and Brown, M. (2007). "Mannahatta: An Ecological First Look at the Manhattan Landscape Prior to Henry Hudson." Northeastern Naturalist, 14(4), 545-570.

[34] Sanderson, E. W. (2009). Mannahatta: A Natural History of New York City. New York: Abrams.

[35] Report of the Tenement House Committee as Authorized by Chapter 479 of the Laws of 1894 (1895). Albany: New York State.

[36] Veiller (1903) "A Statistical Study of New York's Tenement Houses." In: The Tenement House Problem, Volume 1. Eds.: Robert DeForest and Lawrence Veiller. New York: The Macmillen Company.

[37] Veiller, L. (1903a). "Tenement House Reform in New York, 1834-1900." In DeForest and Veiller, Volume 1, 69-118.

[38] Veiller, L. (1903b). "A Statistical Study of New York's Tenement Houses." In DeForest and Veiller, Volume 1, 191-240.

[39] Villarreal, C. (2013). "Where the Other Half Lives: Evidence on the Origin and Persistence of Poor Neighborhoods from New York City 1830-2011." Unpublished manuscript.