Property Emulation:

Consumption Externalities in Renovation Decisions

Ian McGroarty

Haverford College

Department of Economics*

April 2017

Abstract

This paper contributes to the growing literature relating to exogenous social effects in home rehabilitation decisions. The empirical model used in this paper exploits differences between visible and non-visible spillover effects. Specifically, observing differences in the effects of accepted and rejected home improvement loan applications on neighboring renovation demand. This approach controls for non-random neighborhood sorting in addition to endogenous and correlated neighborhood effects. Estimation of this model exploits household level data from the Home Mortgage Disclosure Act that has been aggregated to the neighborhood level. In line with the sociological and economic theory, these results support the hypothesis that demand for home improvement loans in one district increases in response to a positive shift in housing renovation in and adjacent district. My results also show the presence of within neighborhood spillovers. I find that these effects decrease but remain significant with increasing distance. The findings of this paper have important implications for home policy makers and mortgage lending institutions.

*I would like to thank the entire economics department of Haverford College. In particular, my thesis advisers Tim Lambie-Hanson, and Richard Ball. I would also like to give a special thanks to Shannon Mudd and Diana Schoder. Finally, I would like to thank my family for their unending love and support.

Contents

> "*The desire for added comfort and security is present at every stage of the process of accumulation in a modern industrial community: although the standard of sufficiency in these respects is in turn greatly affected by the habit of pecuniary emulation.*" - Thorstein Veblen

1 Introduction

Keeping Up with the Joneses: a popular comic strip created by Arthur Momand. The strip followed the life of the McGinis family, a family of ”social climbers” in their pursuit to match the ever-increasing lifestyle of the neighbors, the Joneses. The comic was popular from the mid-1910s through the late-1930s, an interesting time in the United States, encompassing both the roaring twenties and the great depression. The idea of "Keeping Up with the Joneses" is one popularized by pop culture following a long history of sociological and economic research.

Thorstein Veblen was one of the first academics to popularize the ideas of conspicuous consumption and what he termed "pecuniary emulation". This term refers to the tendency for economic classes to replicate the behavior of those in the class directly above them. Veblen's description of this phenomenon depicts it as inevitable and unyielding: "So long as the comparison is distinctly unfavorable to himself, the normal, average individual will live in chronic dissatisfaction with his present lot; and when he has reached what may be called the normal pecuniary standard of the community, or of his class in the community, this chronic dissatisfaction will give place to a restless straining to place a wider and ever-widening pecuniary interval between himself and this average standard." Veblen (1889). Later writers such as Friedman (1957), speaks out against the presence of this social influence on consumption. The topic of keeping up with the Joneses and, more broadly, consumption externalities, has proven a difficult topic to study empirically, arising various modeling issues.

As income inequality in the United States continues to rise, the relationships in demand between neighborhoods becomes increasingly important. Home rehabilitation decisions contribute to widening economic segregation as socially mobile individuals leave their homes and communities behind. The tendency for wealth to leave areas leads to the decay of infrastructure and further opens the door for gentrification. Better understanding of these issues can aid policy makers in their efforts to combat runaway inequality.

This paper explores the possible consumption externalities that may arise following an increase in the quality of housing infrastructure of one neighborhood. My hypothesis is that an increase in the quality of housing infrastructure in one neighborhood will lead to an increase in demand for home improvement loans in neighboring districts. Using data from the Home Mortgage Disclosure Act (HMDA), I proxy the increase in housing infrastructure with the number of accepted home improvement loans in the district. To measure housing demand, I use the total number of home improvement applications of the district. Following the reasoning of Veblen (1889), an increase in the number of accepted home improvement loans in district A, should precede an increase in the number of home improvement loan applications in district B.

The model presented in this paper leverages this assumption by comparing the effect of rejected loans and accepted loans on neighboring districts. Because rejected loans are not dispersed, these applications do not represent conspicuous consumption. Whereas accepted applications represent home improvement projects that have been carried out. Before an application is approved or denied, an individual must decide to apply. Thus, it is reasonable to expect that the demand side factors that influence the application of accepted and rejected loans are exogenous to the comparison. This implies, that the difference between the effect of accepted and rejected loans of one district on the demand for home improvement loans in another district is unbiased by the endogenous incentives driving renovation decisions. To limit the bias of this assumption I include several district control variables that might influence differences in the overall demand for housing.

My results support the presence of a “Keeping Up with the Joneses Effect" in housing renovation decisions. I find that adjacent tract renovation influences an own tract’s demand for applications, and the relative size of requested loan amounts. These effects diminish with increasing distance, consistent with the “Nearest Neighbors Hypothesis. However, my results show that the spillover effect has an effect beyond the scope of the own tract. My analysis indicate that tracts as far as 10 miles from the own tract will impose a positive exogenous spillover on the own tract. This paper provides has important contributions due to the added specification of an exogenous sorting of the sample. Consequently, my results have important implications for home policy makers, and lending institutions.

I begin by providing some background on the research relating to spillover effects in Section 2. I focus this discussion on the empirical issues that have limited unbiased estimation of spillover effects. In Section 3, I develop my model of renovation decisions. I first, discuss the basic approach applied in the surveyed literature. I follow this by demonstrating how the methodology used in this paper reduces bias in measurement of the exogenous social spillover effect. I then describe the data in Section 4 of the paper. The empirical application to to modeling outline in Section 3 is described and carried out in Section 5. This section is divided into subsections, in which I apply different empirical specifications to the model. The preliminary results are derived using a linear - logarithmic OLS regression analysis comparing the differences in the adjacent tract accepted and rejected effects on the own tract applications. Section 5.3, allows for variation in the size of the spillover effect between near tracts and far tracts. Section 5.4 uses the aggregated total loan amount in the own district as the dependent variable to allow for variation in the dollar for dollar size of the spillover effect. Section 5.5 applies the model to the measurement of lagged own tract spillover effects. The conclusion of this paper discusses potential concerns, expansions, and implications of the results in Section 6.

2 Literature Review

Academics across the social sciences have acknowledged the presence of an exogenous social effect for centuries.Veblen (1889) was one of the first major researchers the bring the phenomena into academic discussion. Despite widespread recognition the task of isolating the effect of the incentive has eluded empirical economist.

2.1 Background

Manski (1993) characterizes three channels through which similarities between the consumption behavior of two groups can be observed.

> “ (a) *endogenous effects,* wherein the propensity of a individual to behave in some way varies with the behavior of the group.
> (b) *exogenous effects,* wherein the propensity of an individual to behave in some way varies with the exogenous characteristics of the group.
> (c) *correlated effects,* wherein individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments." (Manski, 1993, 532)

Galí (1994) argues that rather than bear the cost of decsison making, individuals will imitate the behavior of their peers. Akerlof (1997) identifies the presence of endogenous effects in a non-cooperative game in which individuals were punished for deviating from social norms. Some intution behind Akerlof's game exists in the role of home-owners associations, who impose consequences on lazy home care. Manski (1993), focused on fixing the “reflection problem" implicit in *endogenous effects*", wherein estimations interpret the simultaneous feedback between an individual and a reference group.

Many psychologist have studied *exogenous effects* in small scale purchasing decisions. Chao and Schor (1998) test for conspicuous consumption using cosmetic products and find evidence of a lower correlation between price and quality in products for which a "visibility premium" exists. The visibility premium implies that individuals will emulate the purchasing behavior of a higher class. Thus, their decision is, in part, driven by a jealousy effect. These results are consistent with the findings of Bearden and Etzel (1982), who show that luxury goods are more sensitive to social influence than are necessity goods. Bearden and Etzel also show that there is a stronger incentive to place more weight on the "brand" of a product when the purchase is conspicuous. While these studies, and those like them, typically only represent a small sample and a particular market(s), they demonstrate the nature of human behavior in decision making.

Economists have used varying methodology to identify spillover effects. The first, discussed in the following section, involves a direct approach. Researchers using this methodology attempt to control for all other possible demand factors to isolate the spillover effect.

2.2 Direct Approach

Previous, empirical work focused on an aggregated effect of these three channels. Mayer (1981) find evidence to support the idea that economic returns to investment have a major effect on rehabilitation decisions of landlords. Similarly, Ding et al. (2000) show that residential investment has a positive effect, while spatially limited, on neighboring property values. Case (1991) studies this empirically using a consumer demand model. Her results maintain that, holding price fixed, individual demand increases with the mean demand of a reference group. While, these results are indicative of spillovers, Case and her contemporaries are unable to separate the three spillover channels. In more recent years, economists have adapted their models to bridge this gap.

The burden of empirical analysis is to separate these effects. If one hopes to measure the consumption externalities that are captured by the *exogenous effects*, one must also identify the *endogenous effects* and *correlated effects*. Yannis Ioannides and Jeffrey Zabel have published a significant portion of the research surrounding spillover effects. Between 2001 and 2012 they produced a number of works building on their own empirical models to enhance the robustness of their results. In each of these studies they make use of data from the American Housing Survey and census data to determine the impact of neighborhood renovation expenditure on an individual's own renovation expenditure. Their results can be interpreted as the elasticity of housing demand with respect to the average of neighbors' housing demand. The authors acknowledge, despite their robust models which conclusively determine the presence of endogenous effects, they are unable to separate the exogeneity of neighbors' housing demand. In a recent addition to their series of papers, Ioannides and Zabel (2008) develop a model of housing structure demand with neighborhood effects and neighborhood choice. Their paper is centered on finding and controlling for evidence of non-random sorting into neighborhoods, addressing the key problem in Manski's *correlated effects*. Their results support the hypothesis that individuals prefer to live near others like themselves. This implies that there exists some commonalities between neighbors that might drive similar consumption habits. By controlling for this effect they are able to obtain an unbiased estimate of the elasticity of housing demand with respect to the mean of neighbor's demand. This paper follows two others in which they find and control for *endogenous effects* using a combination of instrumental variables and pricing models. (Ioannides, 2002) (Ioannides and Zabel, 2003). A fourth paper, Ioannides (2010) focuses on spacial econometrics to separate the *contextual effects* that drive *correlated effects*. The reasoning behind this approach follows closely to the nearest neighbors' hypothesis. The spillover from further neighbor's will be uncorrelataed with contrextual effects. This seems a reasonable assumption since further neighbors face different environments. However, this paper is limited in its inclusion of neighborhood characteristics that may lead to overestimation of spatial interactions. For example, the neighborhood sorting conditions outlined in Ioannides and Zabel (2008). In general, the results of these papers limited in the reproduciblity and applications as they rely on comprehensive data to control for omitted variable bias in estimating housing structure demand. In the absence of very comprehensive control variables, it becomes extraordinarily difficult to validate a social influence on consumption by observing that a group of neighbors purchase similar baskets of goods. Thus, the remainder of the literature has focused on using control groups and quasi-experimental data in their attempts to isolate the exogenous effects.

2.3 Experimental Approach

Helms (2012) follows the spacial model of Ioannides (2010) closely, adding a more robust analysis with neighborhood controls using data from the Chicago Department of Buildings and confidential census data. The model relies heavily on the use of a spatial lag weighting the neighbors' choice decisions based on distance rather than time. This approach enables more freedom in terms of measuring the effect by using the correlation between distance and the visibility of renovation. Grinblatt et al. (2008) employ a similar approach in their analysis of car purchase spillovers. This approach, coined the nearest neighbor's hypothesis, relies on controls to isolate exogenous effects from the correlation effects resulting from neighborhood sorting. Their results show a very steep dependence on time and distance in the proliferation of social influence. Beyond the nearest 10 neighbors and after 10 days there is little to no evidence of any consumption externalities.[1] Other research has identified exogenous effects using quasi-experimental data. One major work identifies the presence of a social spillover resulting from an exogenous shock resulting from a lottery. Kuhn et al. (2011) take advantage of the Dutch Postcode Lottery, which allocates prize winnings to participants in a randomly chosen postcode. By comparing the participants and the non-participants, the authors are

[1] I suspect that this is partially due to the nature of cars, which are less visible and more immediately shocking than are housing repairs.

able to separate the effects of an exogenous income shock. They find evidence for the effects of lottery prizes on neighbors of winners in car purchase decisions. However, their estimates do not explicitly distinguish between endogenous and exogenous social effects.

In this paper, I further the attempt to isolate the exogenous social effect in housing renovation decisions. Using application data from the Home Mortgage Disclosure Act, I compare the effects of successful applications and non-successful applications on neighboring rehabilitation demand. This paper's renovation model expands on the nearest neighbor methodology for controlling for unobservable effects. Rather than using distance as a categorization for a control group, I use rejected loans. Comparing the volume of rejected loans to the volume of accepted loans eliminates the bias of omitted variables and non-random sorting. The analysis consists of a series of pair regressions, with the total number of home improvement loan applications in an "own" tract in a given year serving as the dependent variable.

3 Methodology

The core methodology present in the majority of the literature centers on the function for consumer demand. Own demand for housing renovation at time t is a function of some neighborhood characteristics of the individual and some spillover effect of adjacent tract j.[2]

$$Applications_{i,t} = F(characteristics_i), G(spillovers_j) \tag{1}$$

Ioannides and Zabel, tease out the derivatives of F by controlling for the $characteristics_i$. This approach requires a mass of information and complex modeling of structural housing demand functions. The inevitable omitted variables may be correlated with the error term causing bias in the measurement of the spillover effect. Hence, the model makes a distinction

[2] Different papers have used different lags on *characteristics* and *spillovers*, to preserve generality I do not include time subscripts in the theoretical modeling.

between observable and unobservable characteristics.

$$Applications_{i,t} = H(observable_i), K(unobservable_i), L(spillovers_j) \tag{2}$$

While these observable controls reduce some bias, the direct approach encounters significant difficulties. (Case et al., 2011). As one author pointed out, advertising in news articles can have significant impacts on community decision making. The unobservable characteristics lead to bias insofar as they are correlated with spillovers. While the direct approach models can serve to identify the clustering of decisions around neighborhood, the correlation between unobservable characteristics and spillovers limits the separation of exogenous, contextual, and endogenous neighborhood effects. For the purposes of this paper, I am interested in identifying an endogenous social spillover across census tracts.

One popular method of isolating exogenous neighborhood effects is known as the nearest neighbor's hypothesis. These models distinguish between conspicuous and inconspicuous consumption, in an attempt to classify a pseudo control group. The nearest neighbor's hypothesis exploits differences in the visibility of "near" and "far" purchasing behavior.

$$Applications_{i,t} = H(observable_i), K(unobservable_i), M(visible_j), N(invisible_j) \tag{3}$$

For example, an individual is much more likely to witness a new Ferrari cruising around the neighborhood if it belongs to their next-door neighbor, than if it belongs to someone across town. The key assumption underlying this strategy is that it is possible to implicitly net out unobservable characteristics using the non-visible control group. Grinblatt et al. (2008), employs this methodology to identify exogenous spillovers in car purchase decisions. In this paper, the authors show that the difference between the effect of conspicuous spillovers and that of inconspicuous spillovers, (M-N), is uncorrelated with the unobservable characteristics. Thus, the difference in the coefficients is unbiased by unobservable characteristics. However, this approach falls short in controlling for the endogenous location effect. The approach is inherently limited by factors such as nonrandom residential sorting. If individuals tend to live near others who implicitly have similar characteristics, the resulting correlation in decision making will lead to upward bias in the measurement of spillovers. If differences in an omitted variable between the visible and invisible groups are correlated with the purchase decision, and the variable lies outside the reach of the controls, the model will yield biased estimates. In the model that follows, I eliminate this bias by using the exogenous decision of approval, rather than the endogenous decision to purchase.

3.1 Adjacent Tract Spillovers

Following the theory of the nearest neighbor's hypothesis; the coefficients on the observed and the unobserved characteristics, absent spillovers, should be equal. (Grinblatt et al., 2008) (Helms, 2003). From this, I define;

Assumption 1: Absent additional spillover effects, the effect of the Accepted Applications should be equal to the effect of the Rejected Applications.

This assumption is a direct product of the nearest neighbor's hypothesis which uses non-visible purchases (of distant neighbors) as "an instrument for omitted control variables that might generate spurious inferences about visible purchases (near neighbor influence)." (Grinblatt et al., 2008, 735). The comparison between accepted loans and rejected loans is perhaps more robust between near and far neighbors. As Ioannides and Zabel (2003) point out, neighbors tend to sort themselves into neighborhoods with individuals with similar characteristics to their own. Comparing Accepted and Rejected loans, controls for this issue simply because it compares individuals within a reference group. Since the sample used in this paper is restricted to applications, this assumption is unbiased by neighborhood effects (since everyone in the sample has chosen to apply). This assumption has an important qualification:

Assumption 2: The unobservable incentives to apply for a loan are equally correlated with

Accepted and Rejected applications.

This implication may be biased if lower credit borrowers, who are more likely to be rejected, are more vulnerable to certain signals that influence their decision making. Concerns for this implication are limited since our analysis focuses on the tract level rather than the parcel level. These differences will be much easier to control for once aggregated to the tract.[3] Similarly, I make an assumption about the relationship between the exogenous spillover effects and conspicuous consumption:

Assumption 3: The effect of the Rejected loans is uncorrelated with the exogenous social spillover. The effect of the Accepted loans is highly correlated with the exogenous social spillover.

Assumption 3, follows from the idea that, because the purchases are non-visible, they do not have an effect. The home improvement projects for which the individual applied for the loan are never carried out. Since they are never carried out, their neighbors can not get jealous of the change. So problems may arise for this assumption if it individuals are getting their loans from other places. Hence, the results of this paper may not be generalizable to communities who renovate their homes without applying for a home improvement loan.

To illustrate how these assumptions interact, I linearize equation (2).

$$Applications_i = \alpha_o X_o + \alpha_u X_u + \alpha_\rho \rho$$

With $\alpha_o X_o$ the observable characteristics, $\alpha_u X_u$ the unobservable characteristics, and $\alpha_\rho \rho$ the spillover effect. Since unobservable characteristics, such as neighborhood sorting, will be correlated with both accepted and rejected applications, their effect on $Applications_j$ can

[3] More on the difference between tract and parcel level analysis are discussed below.

be proxied as the observable characteristic and modeled:[4]

$$Applications_i = \alpha_0 Accepted_j + \alpha_u Xu + \alpha_\rho \rho \tag{4}$$

$$Applications_i = \alpha_1 Rejected_j + \alpha_u Xu + \alpha_\rho \rho \tag{5}$$

However, since $\alpha_u X_u$ and $\alpha_\rho \rho$ are not specified in the regression, α_o will be biased. To understand how α_o will be biased, recall that both $Accepted_j$ and $Rejected_j$ are correlated with the unobservables $\alpha_u X_u$. I apply the model detailed in (Wonnacott and Wonnacott, 1979), for the solution to "Inconsistency of OLS When e and X are Correlated". The method is a modified instrumental variables approach, where X_u is correlated with $Accepted_j$ and $Rejected_j$ but uncorrelated with ρ:

$$X_u = c_0 + \gamma_1 Accepted_j$$

$$X_u = c_1 + \gamma_2 Rejected_j$$

Substituting these equations into Equation 4 and Equation 5 respectively:

$$Applications_i = (\alpha_0 + \alpha_u \gamma_1) Accepted_j + \alpha_u c_0 + \alpha_\rho \rho \tag{6}$$

$$Applications_i = (\alpha_1 + \alpha_u \gamma_2) Rejected_j + \alpha_u c_1 + \alpha_\rho \rho \tag{7}$$

$(\alpha_0 + \alpha_u \gamma_{1,2})$ is thus the bias on α_o created by unobservable variables uncorrelated with the social spillovers but correlated with $Accepted_j$ and $Rejected_j$. Differences in γ_1 and γ_2 are driven by unobservable characteristics likely to drive differences in a lenders decision. For example, looking at individuals $\gamma_1 - \gamma_2$ may reflect differences in credit scores. However, because this analysis focuses on census tracts, the differences are more likely going to be driven by different tract characteristics such as median income and employment. Additionally,

[4] With a limited set of control variables, it is clear that the tendency for accepted and rejected loans to move in tandem, may raise colinearity concerns. Using two separate regressions allows for an accurate representation of the biased coefficients, while avoiding these concerns.

market trends will influence possible differences in γ_1 and γ_2.

Looking at equations 6 and 7, it is clear that the error term $\alpha_\rho \rho$ will also be correlated with the observables, $Accepted_j$ and $Rejected_j$. Once again applying the modified instrumental variables technique:

$$\rho = c_3 + \delta_1 Accepted_j$$

$$\rho = c_4 + \delta_2 Rejected_j$$

Applying this pair to Equation 6 and Equation 7 respectively yields the system of equations:

$$Applications_i = (\alpha_0 + \alpha_u \gamma_1 + \alpha_\rho \delta_1) Accepted_j + \alpha_u c_0 + \alpha_\rho c_3 + \alpha_\rho \rho \tag{8}$$

$$Applications_i = (\alpha_1 + \alpha_u \gamma_2 + \alpha_\rho \delta_2) Rejected_j + \alpha_u c_1 + \alpha_\rho c_4 + \alpha_\rho \rho \tag{9}$$

To isolate the spillover, take the difference between Equation 8 and Equation 9:

$$\begin{aligned} &(\alpha_0 + \alpha_u \gamma_1 + \alpha_\rho \delta_1) Accepted - (\alpha_1 + \alpha_u \gamma_4 + \alpha_\rho \delta_2) Rejected \\ &\quad + [\alpha_u c_0 + \alpha_\rho c_3 + \alpha_\rho \rho - (\alpha_u c_1 + \alpha_\rho c_4 + \alpha_\rho \rho)] \end{aligned} \tag{10}$$

The term in brackets is the error differential error term, clustered around zero.[5] The differ-

[5] This will be explored later in the paper. In Section 5.1 Residual Justification.

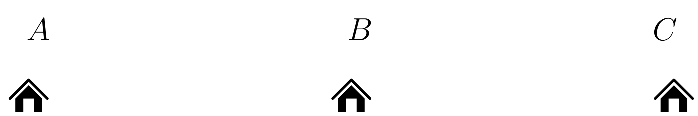
ence in the coefficients shows how to isolate the spillover effect:

$$
\begin{aligned}
&(\alpha_0 + \alpha_u\gamma_1 + \alpha_\rho\delta_1) - (\alpha_1 + \alpha_u\gamma_2 + \alpha_\rho\delta_2) && \\
&\alpha_0 + \alpha_u\gamma_1 + \alpha_\rho\delta_1 - \alpha_1 - \alpha_u\gamma_2 - \alpha_\rho\delta_2 && \\
&\alpha_0 + \alpha_u\gamma_1 + \alpha_\rho\delta_1 - \alpha_0 - \alpha_u\gamma_2 - \alpha_\rho\delta_2 && \text{By Assumption 1: } \alpha_0 = \alpha_1 \\
&\alpha_u\gamma_1 + \alpha_\rho\delta_1 - \alpha_u\gamma_1 - \alpha_\rho\delta_2 && \text{By Assumtion 2: } \gamma_1 = \gamma_2 \\
&\alpha_\rho\delta_1 - \alpha_\rho\delta_2 && \text{By Assumption 3: } \alpha_\rho\delta_2 = 0 \\
&\alpha_\rho\delta_1 &&
\end{aligned}
$$

Hence, the remaining term is Accepted's effect on ρ. Thus, taking the difference in the coefficients on *Accepted* and *Rejected*, leaves only the exogenous social spillover effect.

3.2 Intuition Through Example

To consider the relevance of this model in comparison to spatially lagged models, consider the following: Take three adjacent districts A,B, and C. Let A be the own district and B,C be the adjacent districts of interest, such that B is one mile from A and C is two miles from A.



Define B to have a very low unemployment rate, and C to have a very high unemployment rate. A will have an unemployment rate similar to that of B, also say that the residents of A and B are employed in the same industry, say the air conditioning industry. Due to global warming, there is an extremely hot year, and the air conditioning industry is booming. The increase in job security, leads A and B to embark on home renovation projects at a similar time, and yet C does not feel this shock. If a spatially lagged model cannot control for feelings of job security, then it will likely classify this joint increase as a social spillover, since B influences A but C does not. Thus, a bias arises since the omitted variable, job security, is correlated with the decision to engage in renovation activities.

By using the difference in accepted and rejected loans in B and C, this model looks for differences between someone who receives a loan and someone who does not. Because the sample only includes individuals who decided to engage in renovation activities, people who applied for a loan, any endogenous decision to renovate can be ignored. Since, the air conditioning boom is going to influence a person's decision to apply, not necessarily whether they get accepted. Differences in the effects of the air conditioning boom on the likelihood of acceptance will be captured in other variables such as income.[6] This follows from Assumption 1; since the neighborhood effects will influence individuals within B the same regardless of whether or not their loan was accepted. Using controls to ensure that the lender's decision to either approve or deny an application is exogenous, the model can accurately control for endogenous neighborhood effects.

To see how my model controls for this, examine Figure 1. Researchers using the direct approach, examine the straight line between 'Incentive Adjacent' and 'Incentive Own'. Controlling for all possible other factors that may be correlated with the own tract incentives and the adjacent tract incentives. The nearest neighbor's hypothesis compares near districts and far districts. In the figure it is clear how this method can be biased by endogenous differences between the near districts and far districts. This bias includes, but is not limited to, non-random neighborhood sorting: Since the people choosing the Near path are fundamentally different from the people choosing the Far path. This paper adds the third column comparing accepted and rejected applications. Since the decision to apply is a prerequisite for inclusion in the sample, these endogeneity concerns are nullified. Hence, the primary

[6] This glosses over a major potential problem in of this paper. Due to limited control variables, differences in the likelihood of accepted and rejected may indeed be correlated with some omitted variables and create a bias on the spillover effect. This potential concern is discussed below.

concern for the analysis is controlling for market trends that may create differences between the accepted applicants and the rejected applicants.

4 Data

The Home Mortgage Disclosure Act (HMDA) was passed to combat discrimination in lending by requiring transparency in loan acceptance and denial statistics. The legislation required the collection of data on the lending habits of depository institutions. These data, collected by the Federal Financial Institutions Examination Council (FFIEC), include information on the lending institution, applicant demographics including applicant income, loan characteristics, whether or not a loan is accepted, and if the reason for denial (if the loan was denied). Since its inception in 1975, the legislation on which institutions are required to report HMDA[7] data has changed through subsequent amendments: The attributes which determine whether an institution is required to submit HMDA data is based on the lender's size, whether or not it participates in residential mortgage lending, and whether or not is has a home or branch of the office in a metropolitan statistical area (MSA), among other features. [8]

For this report, I conduct my analysis on a small sample from the original data file retrieved from the public the data warehouse hosted by the Consumer Financial Protection Bureau. This file contains home improvement loan applications over a four-year period 2010 – 2013 for homes located in Pennsylvania. The raw dataset contains 28,816,138 observations and 47 variables. Unlike the confidential HMDA data governed by the legislation, the public data

[7] For more information on the regulation on HMDA institutions, and the types of institutions covered, see (Avery et al., 2007), and (Witowski 2016)

[8] The public HMDA data differs from the confidential HMDA data in two important ways: The confidential data includes unique loan identifiers that can be used to merge with other confidential datasets, the day of the application date and the action date. The public data on the other hand, only includes application and action year rather than the day, and does not include the loan number, though each loan has a non-confidential unique identifier.

classifies loan by an identification year, a census tract, and a unique identifier. A census tract typically represents approximately 4,000 people and attempts to create geographic sections or mostly homogeneous people in respect to socio-economic status, population size, and living standards. The identification year is the year in which the application was submitted and does not represent the date of the action taken by the depository institution. The data includes information on the borrower, lender, the property, the location, and the loan itself. The borrower identifications consists of race, ethnicity, sex, and income. The loan data include; year of application submission, action type (whether or not the loan was accepted), the lien type, amount, purpose, the type of application such as if it was pre-approved, the reason for denial, and the rate spread. The geographic information available describes the census tract, county, MSA, the median income of the tract, the population of the tract, the demographic spread of the tract, the number of housing units, the number of owner occupied houses, and the number of vacant units. Finally, data on the property include, location, the occupancy status, the type such as multifamily dwelling or manufactured home.

The final data in my analysis include only applications filed for home improvement, dropping all applications for home purchase and refinancing. I only include homes that are owner occupied and homes which are the primary residence of the owner. This is so that the home owners are the ones who experience the social influence so including buildings and renters is will create downward bias on the results. Similarly, I exclude manufactured housing which only represents a very small portion of the observations. I include a loan size specification of a minimum limit of $ 5,000 and a maximum of $200,000. This is to exclude very small repairs that may not be as significantly visible to the community. The frequency distribution of the size of the loans in the sample is displayed in Figure: 2. In this figure it is clear that this limitation does not exclude many loans. Finally, I remove all observations for which the respondent id or census tract number is unavailable.

To obtain the geographic information on tracts, I use the NBER census tract distance database. This data set records all census tracts and contains observations for each adjacent tract that is within 25 miles of the "own" tract. For each tract pairing the distance between the centers of the adjacent tract to the own tract is recorded. Pairing this distance data with the HMDA data once, gives me the own tract information with the adjacent tract pairs. I then merge the data again in order to gain the HMDA tract information for each of the adjacent tracts.

To conduct my analysis, I collapse the data on the own tract. The number of accepted and rejected loans is summed to include the total number of accepted and rejected loans in all of the adjacent tracts. The resulting panel dataset contains one observation per tract per year, and the (weighted) sum of the number of accepted and rejected loans for all adjacent tracts.

5 Empirical Estimation

Table 1 presents summary statistics on these variables. With in each tract, between 8-9 applications are rejected, and between 11-12 applications are accepted each year. 'Accepted Adj' and 'Rejected Adj', represent sum of the the corresponding applications for neighboring tracts within 25 miles. With an average 100 adjacent tracts within range, each own tract has experiences 1,175 visible accepted loans and 1,445 rejected applications. For tracts within 10 miles, there are 769 total acceptances and 10,50 total rejections. Each own tract has, on average, 4 neighboring tracts within 10 miles.

I begin my analysis with a discussion of the unique specifications made for each model. I conclude each subsection with a discussion of the results.

5.1 Preliminary

The *Accepted* and *Rejected* variables are taken as thus sum total of the corresponding applications within each tract for each year. The main dependent variable, $Applications_{i,t}$ is defined to be the total number of applications, accepted and rejected, in the own tract, in the observation year. The key independent variables, $Accepted_{j,t-1}$ and $Rejected_{j,t-1}$, are the total applications, disaggregated by accepted and rejected applications, summed over the adjacent tract in the lagged observation year. [9] The resulting dataset has one observation per tract, per year.

Following the approach outlined in the methodology section, I use a series of regression pairs to uncover the role of spillovers in the market for home improvement loans. As a final specification, I take the log of $Accepted_{j,t-1}$ and $Rejected_{j,t-1}$. The logged values help not only with interpretation, but also serve as a way of controlling for the skewed nature of these variables. Figure 2 shows that many tracts have 0 accepted applications, while fewer have between 1-5, and still fewer have more than 10. To formulate the empirical estimation of my model, to a baseline regression, I apply OLS modeling to equations 8 and 9:

$$Applications_{i,t} = \alpha_0 + \beta_A \sum_j (Accepted_{j,t-1}) + \gamma X + \epsilon_A \tag{11}$$

$$Applications_{i,t} = \alpha_0 + \beta_R \sum_j (Rejected_{j,t-1}) + \gamma X + \epsilon_R \tag{12}$$

Where γX represents a series of control variables. My approach implies that the exogenous social spillover effect can be determined by taking the difference between the coefficients on $Accepted_{j,t-1}$ and $Rejected_{j,t-1}$. Thus, my analysis will serve to test the following:

[9] The use of lagged variables is to allow the application to go through the process, for the renovation activity to take place, and for the jealousy factor to take effect. One can imagine an application submitted in March, approved in April. The actual project is carried out through the summer. And the visibility effects are felt through the early winter months, to result in an application in the early months of the following year.

$$\text{Null Hypothesis } H_0: \ \beta_A - \beta_R = 0$$

$$\text{Alternative Hypothesis } H_A: \ \beta_A - \beta_R \neq 0$$

Assuming control variables adequately account for endogenous differences in β_A and β_R, then $\beta_A - \beta_R = \rho$ with ρ representing the spillover effect. ρ can thereby be interpreted as the exogenous social influence of an increase in housing quality, as measured by the volume accepted home improvement loans, in neighboring districts on demand for housing renovation in the own tract.

Columns 1-4 of Table 2: Primary Regressions, presents the results of the baseline analysis presented in equations 11 and 12. Column 2 shows that a one percent increase in accepted loans in the adjacent districts is associated with an additional 1.865 applications in the own district. Column 2 presents the associated rejected pair, and finds the same increase in rejected loans is associated with an extra 1.204 applications in the own district. At the bottom of the column 2, I calculate the difference between these two coefficients. I then determine the significance of this difference using a standard F-Test. My base line model shows evidence of a significantly positive exogenous social influence effect. In columns 3, and 4, I include the lagged acceptances and rejections from the own tract. Unsurprisingly, these coefficients are positive and significant. This inclusion causes the spillover effect to diminish to 0.267, but the difference maintains significance. The drop in magnitude is likely the of result market and tract trends, that would clearly be correlated with accepted and rejected loans in the adjacent tract. The reflects the correlation between market trends, and differences between the accepted applications and rejected applications. This is further supported in the coefficient on Unemployment. The Unemployment rate is also significant and positive, indicating that more unemployment results in more demand. This surprising result is perhaps indicative of substituting sources of income. As a control for this bias, I include the lagged own tract controls. With these controls, the coefficients on Median Income, Loan Amount, Rate Spread and the Rural dummy variable lose all significance. This furthers the idea that the own tract controls are accounting for some market trends. However, it is difficult to say whether this will account for all potential bias in my results. The number of occupied housing units, and the number of families in the tract remain significant confirming the story that neighborhood characteristics play a role in demand.

As an added specification control, I weight $Accepted_{j,t-1}$ and $Rejected_{i,t-1}$ by the distance between tract j and tract i. This allows for differences in the size of the spillover for further tracts with respect to nearer tracts. The distance weight, $\omega_{i,j}$, indicates the hypothesized relative strength of the spillover effect dependent on the distance. Thus, as distance increases the spillover effect decreases. The resulting regression pair includes δ_j, the inverse of the distance of tract j.

$$Applications_{i,t} = \alpha_0 + \beta_A \sum_j \omega_{i,j}(Accepted_{j,t-1}) + \gamma X + \epsilon_A \tag{13}$$

$$Applications_{i,t} = \alpha_0 + \beta_R \sum_j \omega_{i,j}(Rejected_{j,t-1}) + \gamma X + \epsilon_R \tag{14}$$

For the primary regression analysis, I continue to take the log of the key independent variable, $\Sigma_j \omega_{i,j}(Rejected_{j,t-1})$. ρ in this approach is simply a estimate of the spillover effect robust to distance. This coefficient is *not* an interaction term and therefore does not reflect changes in the spillover effect for varying distances. The results of this analysis are presented in columns 5-6 of Table 2: Primary regressions. Consistent with my expectations, the spillover effect decreases, but remains significant. Note that the coefficients on $Accepted_{j,t-1}$ and $Rejected_{j,t-1}$ are similarly diminished by this specification. The resulting change in each coefficient is due to omitted variables correlated with distance. Since, the spillover effect is expected to be correlated with distance, there is a drop in the magnitude. However, other variables that may be correlated with distance, such as non-random neighborhood sorting, equally effects the accepted and rejected loans but does not effect the magnitude of the spillover. For instance, further away tracts may receive a different newspaper, which shows ads for home renovation projects. The newspaper is correlated with the decision to apply and thus will result in upward bias of our estimates. Therefore, taking away the effects of the newspaper by controlling for tract distance will diminish the estimated spillover. However, since the newspaper effects the decision to apply, rather than whether or not the activity was carried out, it is exogenous to the difference between accepted and rejected applications. I show how the spillover effect varies with distance, in the next section.

5.2 Residual Justification

Recall equation 10 which showed that the the difference in the coefficients, resulted in a difference in the error terms:

$$\alpha_u c_0 + \ \alpha_u c_3 + \alpha_\rho \rho - (\alpha_u c_1 + \alpha_\rho c_4 + \alpha_\rho \rho).$$

If the model is robust, the expected value of this term should be zero. Theoretically, the justification for this is twofold: First, I assume that the unobserved effects are equally correlated with *Accepted* and *Rejected* (Assumption 2). If this assumption holds then $c_0 = c_1$. Second, I assume that *Rejected* is uncorrelated with the spillover effect, ρ (Assumption 3). By Assumption 3: $c_4 = 0$. Finally, if the spillover effect is well represented by its effects on A (follows from Assumption 1), then c_3 is clustered around zero. Hence:

$$
\begin{aligned}
&\alpha_u c_0 + \alpha_u c_3 + \alpha_\rho \rho - (\alpha_u c_1 + \alpha_\rho c_4 + \alpha_\rho \rho) \\
&\alpha_u c_0 + \alpha_u c_3 + \alpha_\rho \rho - \alpha_u c_1 - \alpha_\rho c_4 - \alpha_\rho \rho \\
&\alpha_u c_0 + \alpha_u c_3 - \alpha_u c_1 - \alpha_\rho c_4 \\
&\alpha_u c_0 + \alpha_u c_3 - \alpha_u c_0 - \alpha_\rho c_4 && \text{By Assumption 1} \\
&\alpha_u c_3 - \alpha_\rho c_4 \\
&\alpha_u c_3 - \alpha_\rho c_3 && \text{By Assumption 2} \\
&\alpha_u c_3 \approx 0
\end{aligned}
$$

I show the results of these assumptions in Figure 3. The top two panels show the error terms of the individual regressions from models 2 and 4 in Table 2.[10] Consistent with expectations, the error terms are each severely correlated with $Applications_{j,t-1}$. The bottom panel shows the difference between these two residuals. The plot depicts clustering very close to zero. Thus, I can conclude that my assumptions are robust to residual analysis.

To show how this issue bias's the analysis I include a regression using both accepted and rejected loans, combing models 3 and 4 the resulting estimations are shown in Appendix: Table 6 and the corresponding residual plot, Appendix: Figure 4. Note how the error term remains highly correlated with the independent variable, indicating a biased regression. Hence, I avoid drawing interpretation of spillover effects from single regressions with both accepted and rejected tract effects.

With this residual justification I continue the discussion of my empirical models and their results. In section 5.3 I employ sampling specifications similar in form those used by Helms (2012).

[10]The residual plots from the other models follow identical patterns. The decision not to include all of these graphs was based on the similarities between them, and space.

5.3 Distance

Many papers have made use of spatial lags in their estimation of spillover effects. Helms (2012), weights neighboring renovation expenditure by the distance from the own unit. Following this approach, I weight each adjacent tract by the inverse of the distance to the own tract. To estimate differences in the spillovers between far tracts and near tracts, I limit the sample to include only those tracts with neighboring tracts within 1, 5, and 10 miles. This sample may exclude tracts in very rural areas without close-by neighboring towns which limits generalizations of my results. The sample is thus diminished from 2,370 tracts to 931 tracts. I regress equations 13 and 14 on three different samples; tracts within one mile, tracts further than 1 mile but closer than 5 miles, and tracts between 5 and 10 miles. For this approach, I do not use a linear-logarithmic regression to preserve the the spatial interdependence.

The results of this analysis are presented in Table 3: Distance Sampling Analysis. The difference in the estimates for each of the samples is positive and significant, confirming the presence of an exogenous social spillover in home renovation decisions. The significance and signs on the control variables are consistent with those found in the primary results, indicating a the robustness of the approach. Comparing the spillover effect between the samples confirms the nearest neighbor's hypothesis (Grinblatt et al., 2008). An additional renovation in the adjacent tracts will be associated with 0.0687 more applications in the own tract. This means that for about 11 loans dispersed in adjacent tracts, there is an additional loan application in the own tract. Recall for context that the average number of accepted loans in the adjacent tracts is 1,174. Of these 11 applications, about 2 are resultant of an exogenous social spillover effect. This effect is smaller for tracts further away. Notice, the difference between the spillover effect of the < 1 mile tracts and the $1 - 5$ mile tracts is very large. Compare this change to that between the 1-5 mile sample and the 10+ mile sample. The first increase in distance is very large compared to the second. This is indicative of an exponential spatial interdependence, consistent with Helms (2012). Nearer neighbors are highly visible, while further neighbors are much less visible. The results of this section support the hypothesis that invisibility has exhibits diminishing returns to distance.

5.4 Loan Size

Similar to the spatial interdependence of the spillover effects, one might expect the size of the renovation to have varying effects. It is reasonable to expect that individuals who were already planing a home rehabilitation project may still experience some spillover effect. To test for this effect, I use the total amount of funds requested, as an independent variable. This will allow for variation in the size of the spillover effects. As in the previous models, the independent variable serves as a proxy for the demand for home rehabilitation in the own tract. In this model, the sum of the loan amount for all applications is used rather than the number of loan applications. Note, that this variable reflects applications not necessarily dispersed loans. Taking the total amount of the loans applied for in the own district will allow for a dollar-for-dollar interpretation of the spillover effect. I first use the baseline controls approach, comparing the accepted and rejected loan applications in the adjacent district. I then weight the independent variables by their respective loan amounts. The resulting analysis can be represented by the OLS regression:

$$Loan\ Size_{i,t} = \alpha_0 + \beta_A \sum_j \phi_{j,t-1} Accepted_{j,t-1} + \gamma X + \epsilon_R$$

$$Loan\ Size_{i,t} = \alpha_0 + \beta_A \sum_j \psi_{j,t-1} Rejected_{j,t-1} + \gamma X + \epsilon_R$$

Where $\phi_{j,t-1}$ is the average loan size for the accepted loans and $\psi_{j,t-1}$ is the average loan size for rejected loans in the adjacent tract.[11] As specified. $Funds\ Requested_{i,t}$ is the total

[11] In the base line regression, $\phi_{j,t-1} = \psi_{j,t-1} = 1$

amount, in dollars, requested for home improvement loans in the own tract. The coefficient on these regressions will represent the change in dollars demanded in the own tract, for every additional dollar given in the adjacent tract. In addition to using lagged own tract accepted and rejected applications, I include own tract lagged loan amount.

The results of the loan size analysis are presented in Table 4. Columns 1-2 present the results of the baseline specification. The difference here represents the effect of a one percent increase in the number of accepted loans in the adjacent districts, relative to rejected loans, on the total amount requested by the own district. My results indicate that an a one percent increase in the accepted loans in the adjacent tract leads to \$39,830 additional dollars requested in the own district, of which \$9,380 is attributable to an exogenous social spillover. To place this result into context, recall that the average loan size for my sample is \$51,089. With an average of 19 applications per tract per year, the average amount requested is \$1,019,736. Moreover, a one percent increase in the spillover catalyst represents a large number of accepted loans. More interesting, the results of the weighted analysis, are presented in columns 3-4. Surprisingly, the coefficients have gone negative! Although, $Accepted_{j,t-1}$ is not significantly different from zero, and $Rejected_{j,t-1}$ is only slightly significantly different from zero. It is likely that very small random effects, such as weather, could be responsible for the negative result. This effect would only show up if the coefficients were small enough; which, in this case they are. However, while the coefficients on the two key variables are negative and insignificant, the difference between them is still positive and significant. Thus, I can still reject the null hypothesis at the 0.01 level. The 0.000000437, spillover effect represents the increase in an extra dollar spent in the adjacent districts on the dollars requested in the own district.

5.5 Own Tract Spillovers

While this paper's primary concern was to analyze the spillover effects of adjacent tracts onto an own tract, it is possible to analyze spillover effects within a tract. The literature surveyed in this paper concerns itself primarily with spillover effects within a tract, and thus the presence of own tract spillovers is well documented. Concerns relating to feedback effects and neighborhood sorting have limited the ability to tease out spillover effects between neighbors. The methodology outlined in the previous section shows that, by controlling for potential differences in the effect of accepted and rejected loans on renovation demand, any residual can be attributed to spillover effects. By applying this analysis to lagged own tract acceptances and denials, I am able to eliminate large potential biases in the estimation of own tract spillover effects. Some bias may arise from a stronger correlation between β_A and β_R, for example hysteresis effects may cause an overestimation of the effect of rejected loans with respect to accepted loans. While some bias may persist, the results are still indicative of the presence of a spillover effect. To apply the methodology to own tract spillovers, I simply switch the role of adjacent tracts to control for market and time variance, and use own tract fixed effects, and compare the coefficients on the own tract acceptances and denials. To estimate these effects, I modify the baseline regression pair presented in equations 8 and 9, to use own tracts as independent variables:

$$Applications_{i,t} = \alpha_0 + \beta_A \sum_j (Accepted_{i,t-1}) + \gamma X + \epsilon_A$$

$$Applications_{i,t} = \alpha_0 + \beta_R \sum_j (Rejected_{i,t-1}) + \gamma X + \epsilon_R$$

Rather than own tract lagged controls, I use employ adjacent tract lagged controls, to mediate general market trends. The results of this analysis are presented in Table 5: Own Tract Analysis. As expected the spillover effect is much larger in the own tract, even after controlling for market trends. The spillover effect in this specification, reflects a one percent increase in the number of loans, relative to the number of applications, in the own tract in the previous year on the number of applications in the own tract for the observation year. These results are fairly consistent with the literature.

6 Conclusion

The presence of neighborhood spillover effects has been recognized in the social sciences for centuries. Researchers have used a wide variety of methods to tease out this effect, battling neighborhood sorting and various reflection biases. Along the way, these academics have identified financial as well as social incentives to increase renovation activity. Limitations inherent to data have left many unconvinced of the true nature of spillover effects. This paper takes advantage of the unique HMDA dataset to isolate visible spillovers. I analyzed 6,318 census tracts in Pennsylvania from 2010-2013, and found conclusively that there is a significant difference between visible and non-visible spillovers between neighborhoods. This conclusion is contrary to Friedman (1957) who contends that consumption externalities do not exist at all. Additionally, my results contradict those of Grinblatt et al. (2008) who argue that spillover effects do not extend past the ten nearest unit neighbors.

The results of this paper suggest the presence of consumption externalities, a "Keeping Up with the Joneses" effect on housing renovation decisions. It is therefore reasonable to expect that the effects of a public policy reform aimed at stimulating housing renovation are likely to be spatially multiplied. If a policy's primary purpose is to spur activity (rather than facilitate activity), it would be best to localize the target population in order to allow for the social effects to permeate outward to neighboring districts. Finally, the results have important implications for questions relating to access to finance. The demand for loans is not independent of social incentives as much literature assumes.

The main cause for concern in this paper surrounds omitted variable bias. Controlling for endogenous differences between the accepted group and the rejected group could significantly bias my results. If differences in an omitted variable between the accepted and rejected groups were correlated with the own tract's to apply, it could be the spurious source of my results. I believe that this could be driving my results, and hope that in future studies, with more comprehensive data, this effect is revisited. Another source of interest for this effect would be to study different states. I predict that my results are fairly generalizeable, owing to the diversity in land and in culture of Pennsylvania. Additionally, it may be interesting to apply my model to the time period surrounding the financial crisis. The social incentives during this time would likely be very different give the public interest in the home mortgage market.

One of the distinguishing features of this study is the analysis of tract level effects as opposed to parcel level. This paper is one of very few that have studied this effect. However, applying the model to parcel level data may yield more significant results with greater magnitude. Additionally, this model can be applied to other types of loan based purchases, including, perhaps, credit cards. In relation to this data, it would be extremely useful to include a more robust set of controls to ensure that difference in the effects only captures spillovers. Further analysis into the types of districts that exhibit the greatest spillovers would also be fascinating. Using tract characteristics to identify similar tracts may provide evidence of pecuniary emulation.

References

George A. Akerlof. Social Distance and Social Decisions. *Econometrica*, 65(5):1005–1027, 1997. ISSN 0012-9682. doi: 10.2307/2171877. URL http://www.jstor.org/stable/2171877.

Robert Avery, Kenneth Brevoort, and Glenn Canner. Opportunities and Issues in Using HMDA Data. *Journal of Real Estate Research*, 29(4):351–380, January 2007. ISSN 0896-5803. doi: 10.5555/rees.29.4.wn160840825t7077. URL http://aresjournals.org/doi/abs/10.5555/rees.29.4.wn160840825t7077.

William O. Bearden and Michael J. Etzel. Reference Group Influence on Product and Brand Purchase Decisions. *Journal of Consumer Research*, 9(2):183–194, September 1982. ISSN 0093-5301, 1537-5277. doi: 10.1086/208911. URL http://jcr.oxfordjournals.org/content/9/2/183.

Karl E. Case, Ray C. Fair, and Sharon Oster. *Principles of Economics, Student Value Edition*. Prentice Hall, Place of publication not identified, 10 edition edition, January 2011. ISBN 978-0-13-255298-1.

Angela Chao and Juliet B. Schor. Empirical tests of status consumption: Evidence from women's cosmetics. *Journal of Economic Psychology*, 19(1):107–131, February 1998. ISSN 0167-4870. doi: 10.1016/S0167-4870(97)00038-X. URL http://www.sciencedirect.com/science/article/pii/S016748709700038X.

Chengri Ding, Robert Simons, and Esmail Baku. The Effect of Residential Investment on Nearby Property Values: Evidence from Cleveland, Ohio. *Journal of Real Estate Research*, 19(1):23–48, 2000. URL https://ideas.repec.org/a/jre/issued/v19n12000p23-48.html.

Milton Friedman. *A Theory of the Consumption Function*. Martino Fine Books, 1957. ISBN 978-1-61427-812-2.

Jordi Galí. Keeping up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices. *Journal of Money, Credit and Banking*, 26(1):1–8, 1994. ISSN 0022-2879. doi: 10.2307/2078030. URL http://www.jstor.org/stable/2078030.

Mark Grinblatt, Matti Keloharju, and Seppo Ikaheimo. Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors. *The Review of Economics and Statistics*, 90(4):735–753, 2008. URL https://papers.ssrn.com/abstract=995855.

Andrew C. Helms. Understanding gentrification: an empirical analysis of the determinants of urban housing renovation. *Journal of Urban Economics*, 54(3):474–498, November 2003. ISSN 0094-1190. doi: 10.1016/S0094-1190(03)00081-0. URL http://www.sciencedirect.com/science/article/pii/S0094119003000810.

Andrew C. Helms. Keeping up with the Joneses: Neighborhood effects in housing renovation. *Regional Science and Urban Economics*, 42(1–2):303–313, January 2012. ISSN 0166-0462. doi: 10.1016/j.regsciurbeco.2011.07.005. URL http://www.sciencedirect.com/science/article/pii/S0166046211001001.

Yannis M. Ioannides. Residential neighborhood effects. *Regional Science and Urban Economics*, 32(2):145–165, March 2002. ISSN 0166-0462. doi: 10.1016/S0166-0462(01)00082-5. URL http://www.sciencedirect.com/science/article/pii/S0166046201000825.

Yannis M. Ioannides. Neighborhood Effects and Housing. 2010. URL https://ideas.repec.org/p/tuf/tuftec/0747.html.

Yannis M. Ioannides and Jeffrey E. Zabel. Neighbourhood effects and housing demand. *Journal of Applied Econometrics*, 18(5):563–584, September 2003. ISSN 1099-1255. doi: 10.1002/jae.740. URL http://onlinelibrary.wiley.com/doi/10.1002/jae.740/abstract.

Yannis M. Ioannides and Jeffrey E. Zabel. Interactions, neighborhood selection and housing demand. *Journal of Urban Economics*, 63(1):229–252, January 2008. ISSN 0094-1190. doi: 10.1016/j.jue.2007.01.010. URL `http://www.sciencedirect.com/science/article/pii/S0094119007000356`.

Peter Kuhn, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn. The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery. *The American Economic Review*, 101(5):2226–2247, 2011. ISSN 0002-8282. URL `http://www.jstor.org/stable/23045637`.

Charles F. Manski. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531–542, 1993. ISSN 0034-6527. doi: 10.2307/2298123. URL `http://www.jstor.org/stable/2298123`.

Neil S. Mayer. Rehabilitation decisions in rental housing: An empirical analysis. *Journal of Urban Economics*, 10(1):76–94, July 1981. ISSN 0094-1190. doi: 10.1016/0094-1190(81)90024-3. URL `http://www.sciencedirect.com/science/article/pii/0094119081900243`.

Thorstein Veblen. "Pecuniary Emulation". In *The Theory of the Leisure Class*. Courier Corporation, Chapter 2, 1889. ISBN 978-0-486-28062-2.

Ronald Wonnacott and Thomas Wonnacott. *Econometrics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, 2 edition, 1979.

7 Tables and Figures

	count	mean	sum	sd	min	max
(mean) Unemployment	3568	8.077158	28819.3	1.426757	5.3	10.9
(mean) Distance	3568	9.221466	32902.19	1.626355	4.322947	14.8007
(mean) Median Income	3568	65.81048	234811.8	8.679036	49.9	91.9
(mean) Owner Occupied	3568	1290.373	4604051	535.099	49	3262
(mean) Family	3568	1749.741	6243077	620.892	248	4173
(mean) Loan Amount	3568	51089.12	1.82e+08	27469.67	7333.333	199500
(mean) Minority Population (pct)	3568	15.6797	55945.16	25.12643	.37	99.69
(mean) Rural	3568	.0908072	324	.287375	0	1
(mean) Rate Spread	3568	3.832638	13674.85	1.872733	1.5	14
(mean) Population	3568	4411.104	1.57e+07	1598.503	718	11497
(mean) Applications	3568	19.96216	71225	10.09117	0	68
(mean) Rejected Own	3568	8.5412	30475	5.875351	0	53
(mean) Accepted Own	3568	11.42096	40750	6.694832	0	44
(sum) Accepted Adj	3568	1174.867	4191925	1061.275	0	3792
(sum) Rejected Adj	3568	1444.795	5155029	1826.171	0	6536
(sum) Accepted Adj(10 Miles)	3200	769.2937	2548515	730.3791	0	2629
(sum) Rejected Adj(10 Miles)	3200	1050.466	3844904	1412.16	0	5400

Table 1: Summary Statistics

$Applications_{i,t}$	(1) Baseline	(2)	(3) Controlled	(4)	(5) Distance	(6)
$Accepted_{j,t-1}$	1.865***		1.081***		0.911***	
	(9.66)		(6.43)		(5.99)	
$Rejected_{j,t-1}$		1.204***		0.805***		0.683***
		(7.39)		(5.68)		(5.17)
$Accepted_{i,t-1}$			0.573***	0.589***	0.575***	0.589***
			(21.46)	(21.96)	(21.51)	(21.93)
$Rejected_{i,t-1}$			0.437***	0.434***	0.440***	0.436***
			(14.95)	(14.66)	(15.02)	(14.70)
Unemployment	1.297***	1.252***	0.597***	0.553***	0.571***	0.540***
	(11.00)	(10.45)	(5.73)	(5.27)	(5.47)	(5.12)
Median Income	-0.138***	-0.134***	-0.0584**	-0.0643**	-0.0469*	-0.0501*
	(-5.59)	(-5.05)	(-2.73)	(-2.84)	(-2.24)	(-2.29)
Loan Amount	-0.0506***	-0.0520***	-0.0358***	-0.0369***	-0.0337**	-0.0346**
	(-3.99)	(-4.07)	(-3.30)	(-3.38)	(-3.09)	(-3.18)
Occupied Units	0.00812***	0.00841***	0.00356***	0.00356***	0.00380***	0.00377***
	(11.01)	(11.32)	(5.48)	(5.45)	(5.88)	(5.80)
Families	0.00413***	0.00385***	0.00219***	0.00208***	0.00214***	0.00202***
	(7.08)	(6.54)	(4.33)	(4.09)	(4.22)	(3.97)
Minority Population	0.00336	0.00644	0.00144	0.00265	0.000471	0.000522
	(0.39)	(0.74)	(0.19)	(0.34)	(0.06)	(0.07)
Population	-0.000227	-0.000220	-0.000104	-0.0000886	-0.000149	-0.000121
	(-0.99)	(-0.95)	(-0.53)	(-0.45)	(-0.76)	(-0.61)
Rate Spread	-0.0967	-0.0897	-0.0760	-0.0756	-0.0703	-0.0699
	(-1.22)	(-1.12)	(-1.12)	(-1.11)	(-1.04)	(-1.03)
Rural	0.908	0.388	0.657	0.472	0.762	0.558
	(1.54)	(0.66)	(1.30)	(0.94)	(1.49)	(1.09)
_cons	-9.280***	-4.626*	-4.743**	-2.063	-0.132	0.896
	(-5.06)	(-2.56)	(-2.89)	(-1.30)	(-0.08)	(0.53)
Difference		0.661***		0.276***		0.228***
χ^2		(106.26)		(28.84)		(27.53)
N	2370	2370	2370	2370	2370	2370
adj. R^2	0.486	0.478	0.625	0.623	0.624	0.623

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Primary Regressions

$Applications_{i,t}$	1 Mile		5 Miles		10 Mile	
$Accepted_{j,t-1}$	0.0687		0.00608		0.00210	
	(1.38)		(1.49)		(0.97)	
$Rejected_{j,t-1}$		0.0296		0.000400		0.000295
		(0.87)		(0.16)		(0.19)
$Accepted_{i,t-1}$	0.314***	0.318***	0.331***	0.327***	0.329***	0.325***
	(5.89)	(5.98)	(6.25)	(6.15)	(6.20)	(6.11)
$Rejected_{i,t-1}$	0.400***	0.403***	0.401***	0.409***	0.404***	0.409***
	(10.30)	(10.31)	(10.47)	(10.63)	(10.50)	(10.57)
Unemployment	0.616***	0.641***	0.619***	0.671***	0.660***	0.684***
	(4.24)	(4.39)	(4.38)	(4.60)	(4.74)	(4.79)
Median Income	-0.0675*	-0.0651*	-0.0718*	-0.0591	-0.0732*	-0.0617
	(-2.19)	(-2.10)	(-2.30)	(-1.87)	(-2.20)	(-1.82)
Loan Amount	-0.0151	-0.0144	-0.0192	-0.0175	-0.0182	-0.0167
	(-0.89)	(-0.85)	(-1.17)	(-1.06)	(-1.10)	(-1.01)
Occupied Units	0.00759***	0.00757***	0.00732***	0.00747***	0.00736***	0.00748***
	(8.21)	(8.18)	(8.01)	(8.20)	(8.00)	(8.12)
Families	0.00244**	0.00249**	0.00257**	0.00255**	0.00260**	0.00260**
	(2.64)	(2.67)	(2.87)	(2.83)	(2.90)	(2.90)
Minority Population	0.0171	0.0166	0.0153	0.0201*	0.0194*	0.0202*
	(1.95)	(1.76)	(1.67)	(2.18)	(2.28)	(2.38)
Population	-0.000203	-0.000229	-0.000216	-0.000254	-0.000229	-0.000269
	(-0.64)	(-0.71)	(-0.70)	(-0.81)	(-0.73)	(-0.86)
Rate Spread	-0.0833	-0.0856	-0.0816	-0.0851	-0.0847	-0.0836
	(-0.89)	(-0.92)	(-0.88)	(-0.92)	(-0.92)	(-0.90)
Rural	0.388	0.333	0.450	0.376	0.358	0.355
	(0.10)	(0.09)	(0.12)	(0.10)	(0.09)	(0.09)
_cons	-0.239	-0.457	0.0473	-0.954	-0.169	-0.939
	(-0.11)	(-0.20)	(0.02)	(-0.41)	(-0.07)	(-0.39)
Difference		0.0391***		0.00568***		0.001805***
χ^2		(2.41)		(9.17)		(6.06)
N	931	931	931	931	931	931
adj. R^2	0.697	0.697	0.702	0.701	0.701	0.701

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Distance Sampling Analysis

$Loan\ Amount_{i,t}$	(1)	(2)	(3)	(4)
	Baseline		Weighted	
$Accepted_{j,t-1}$	39.83***		0.000000513*	
	(5.92)		(2.30)	
$Rejected_{j,t-1}$		30.45***		5.93e-08
		(5.37)		(0.34)
$Accepted_{i,t-1}$	23.67***	24.27***	24.09***	23.94***
	(22.17)	(22.65)	(22.37)	(22.11)
$Rejected_{i,t-1}$	10.29***	10.11***	10.98***	11.46***
	(8.73)	(8.49)	(9.20)	(9.55)
L.Loan Amount	-0.673	-0.691	-0.666	-0.603
	(-1.64)	(-1.68)	(-1.61)	(-1.45)
Unemployment	14.04***	12.34**	13.79**	14.13**
	(3.36)	(2.94)	(3.27)	(3.22)
Median Income	-2.515**	-2.799**	-0.940	-0.297
	(-2.93)	(-3.08)	(-1.12)	(-0.33)
Loan Amount	12.93***	12.90***	12.88***	12.93***
	(29.01)	(28.89)	(28.68)	(28.72)
Occupied Units	0.133***	0.133***	0.150***	0.157***
	(5.12)	(5.07)	(5.74)	(6.02)
Families	0.0876***	0.0842***	0.0616**	0.0556**
	(4.33)	(4.14)	(3.11)	(2.82)
Minority Population	-0.262	-0.227	-0.0743	0.0997
	(-0.84)	(-0.72)	(-0.23)	(0.31)
Population	-0.00397	-0.00333	-0.00234	-0.00370
	(-0.51)	(-0.42)	(-0.29)	(-0.47)
Rate Spread	-0.995	-1.011	-0.0640	0.396
	(-0.37)	(-0.37)	(-0.02)	(0.15)
Rural	28.02	21.98	-0.358	-4.727
	(1.39)	(1.09)	(-0.02)	(-0.24)
_cons	-585.6***	-486.4***	-446.9***	-478.6***
	(-8.89)	(-7.59)	(-6.66)	(-6.53)
Difference		9.38***		0.000000437***
χ^2		(19.96)		(23.48)
N	2370	2370	2370	2370
adj. R^2	0.660	0.659	0.656	0.655

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Loan Amount as Dependent

$Applications_{i,t}$	(1)	(2)	(3)	(4)
	Baseline		Distance	
$Accepted_{i,t-1}$	5.033***		4.821***	
	(20.82)		(19.37)	
$Rejected_{i,t-1}$		3.972***		3.789***
		(16.82)		(15.99)
$Accepted_{j,t-1}$			1.771**	5.534***
			(2.76)	(8.72)
$Rejected_{j,t-1}$			-0.207	-3.725***
			(-0.39)	(-7.06)
Unemployment	0.993***	1.012***	0.954***	1.183***
	(8.77)	(8.67)	(8.14)	(10.03)
Median Income	0.0369	-0.0622**	-0.0556*	-0.0547*
	(1.80)	(-2.96)	(-2.17)	(-2.09)
Loan Amount	-0.0517***	-0.0350**	-0.0511***	-0.0307*
	(-4.32)	(-2.82)	(-4.32)	(-2.52)
Occupied Units	0.00571***	0.00764***	0.00466***	0.00688***
	(8.09)	(10.75)	(6.56)	(9.72)
Families	0.00247***	0.00174**	0.00366***	0.00279***
	(4.65)	(3.18)	(6.72)	(4.96)
Minority Population	0.0580***	-0.00273	0.0385***	-0.00924
	(7.17)	(-0.33)	(4.64)	(-1.11)
Population	-0.000243	-0.000145	-0.000241	-0.000243
	(-1.12)	(-0.65)	(-1.13)	(-1.11)
Rate Spread	-0.00208	-0.0860	-0.0625	-0.0957
	(-0.03)	(-1.12)	(-0.85)	(-1.27)
Rural	-0.864	-0.361	0.433	0.920
	(-1.62)	(-0.65)	(0.79)	(1.63)
_cons	-11.59***	-3.195	-15.12***	-17.50***
	(-6.71)	(-1.83)	(-6.45)	(-7.30)
Difference		1.061***		1.032***
χ^2		(14.52)		(14.05)
N	2339	2339	2339	2339
adj. R^2	0.544	0.520	0.557	0.539

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Own Tract Analysis

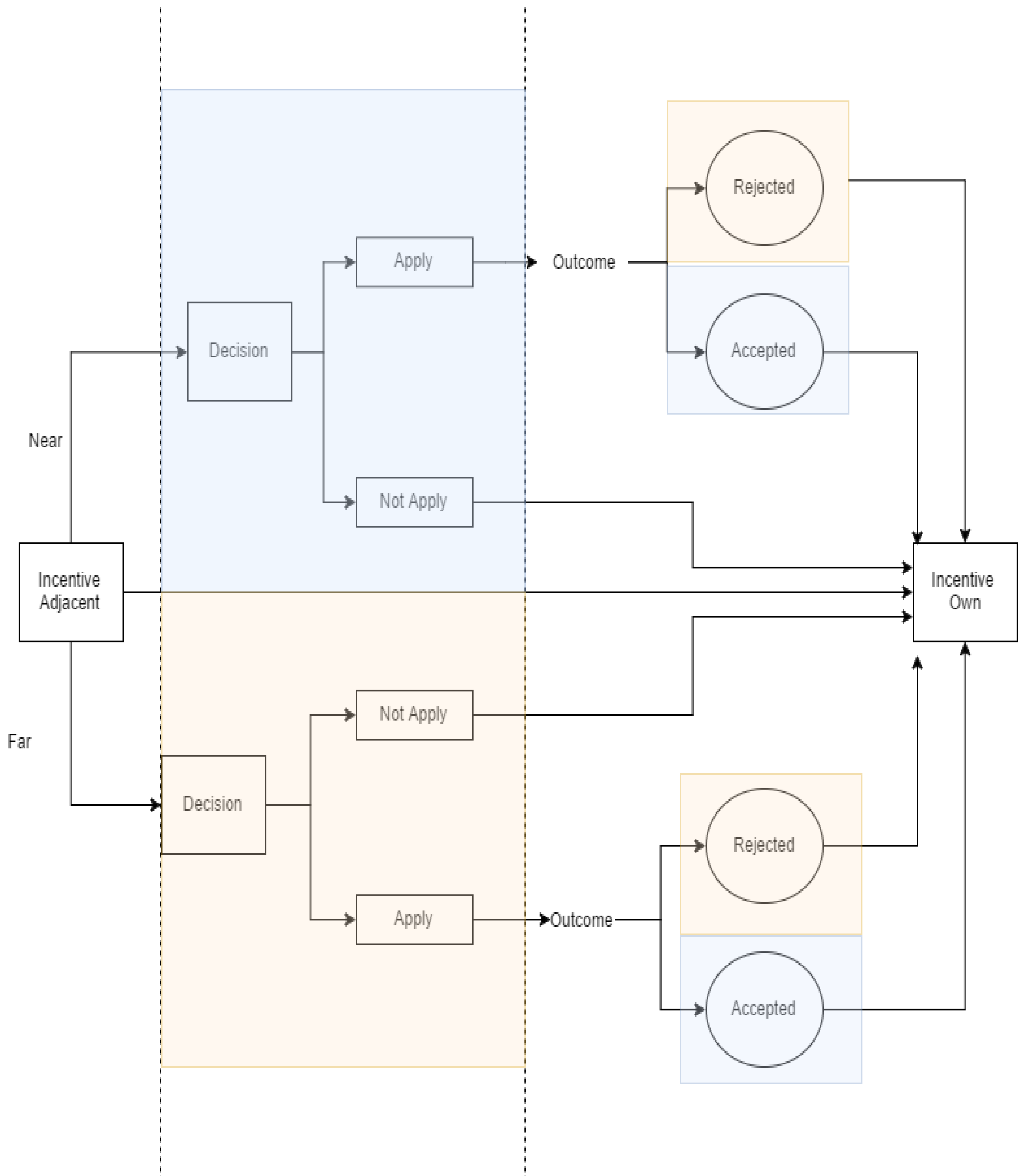


Figure 1: The blue shaded areas represent visible consumption. The orange shaded areas are non-visible.

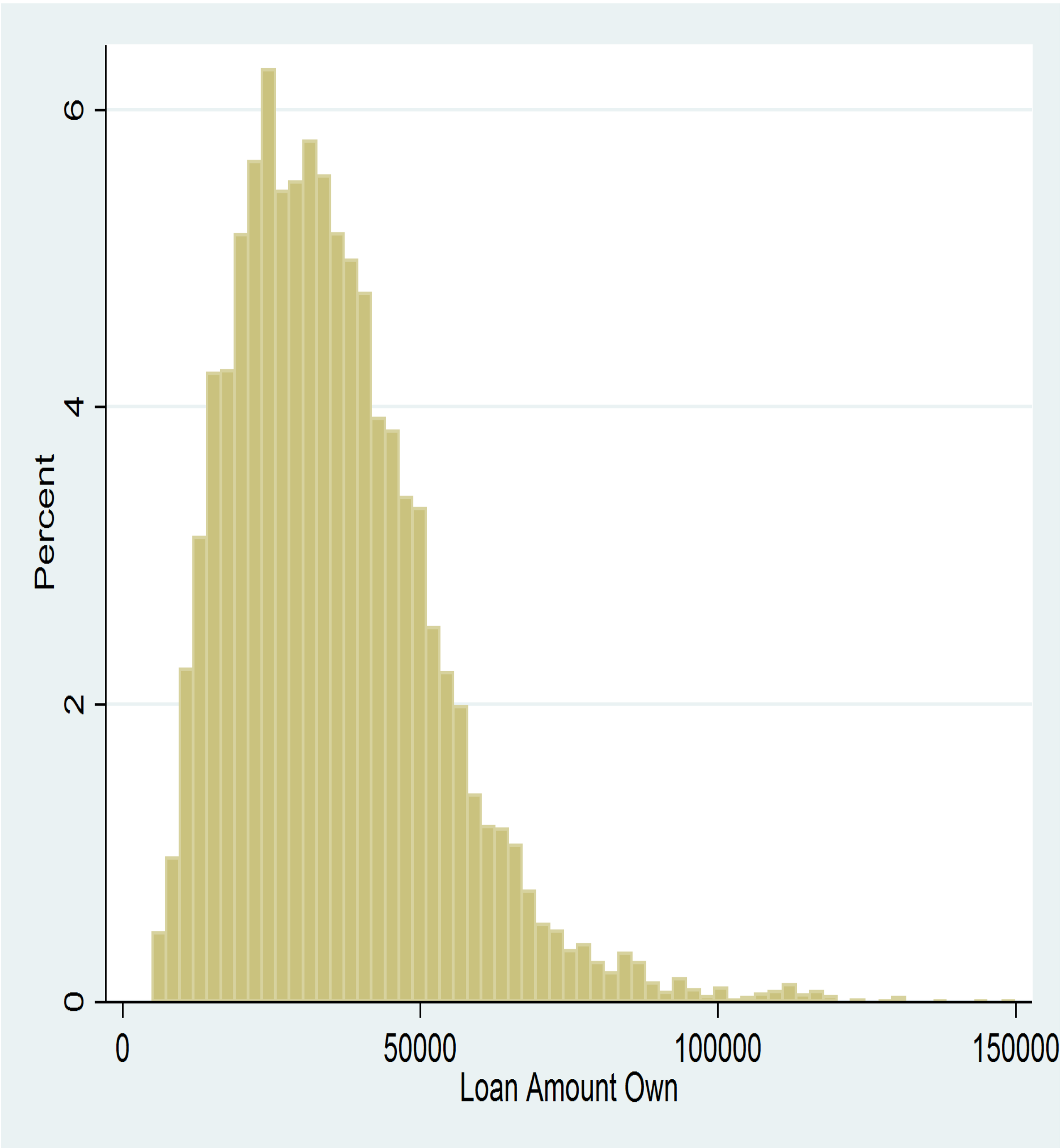


Figure 2: Loan Amount Frequency

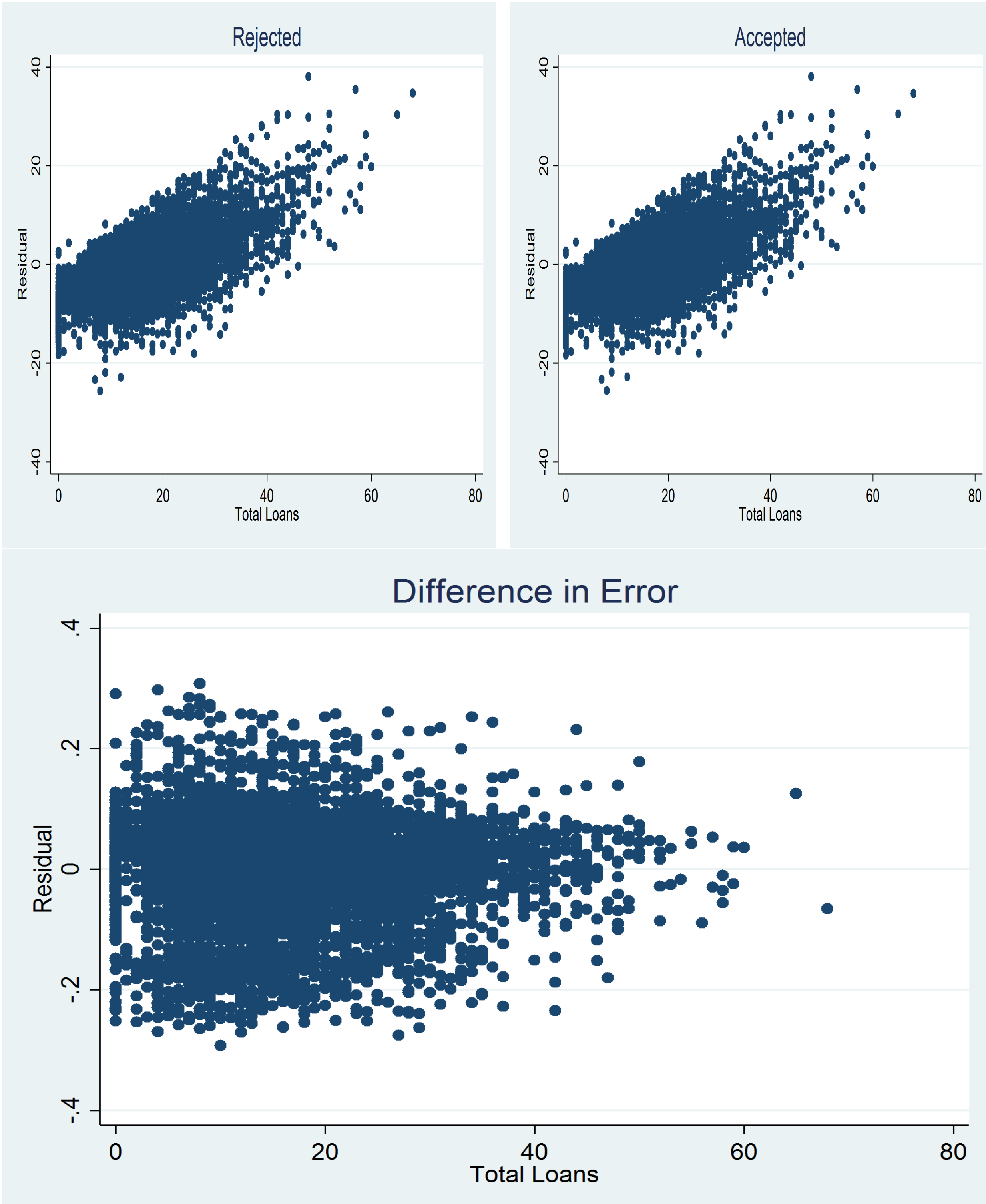


Figure 3: Residual Plots

$Applications_{j,t}$	Controlled
$Accepted_{j,t-1}$	0.00147**
	(3.27)
$Rejected_{j,t-1}$	-0.000417
	(-1.57)
$Rejected_{i,t-1}$	0.488***
	(16.57)
$Accepted_{i,t-1}$	0.595***
	(22.10)
Median Income	-0.0287
	(-1.32)
Occupied Units	0.00328***
	(5.06)
Families	0.00203***
	(4.14)
Loan Amount	-0.0456***
	(-4.20)
Minority Population	0.0134
	(1.68)
Population	-0.0000625
	(-0.32)
Interest Rate	-0.119
	(-1.75)
Rural	0.244
	(0.49)
_cons	4.168**
	(2.91)
Difference	0.001887
χ^2	(0.49)
N	2386
adj. R^2	0.617

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Appendix table

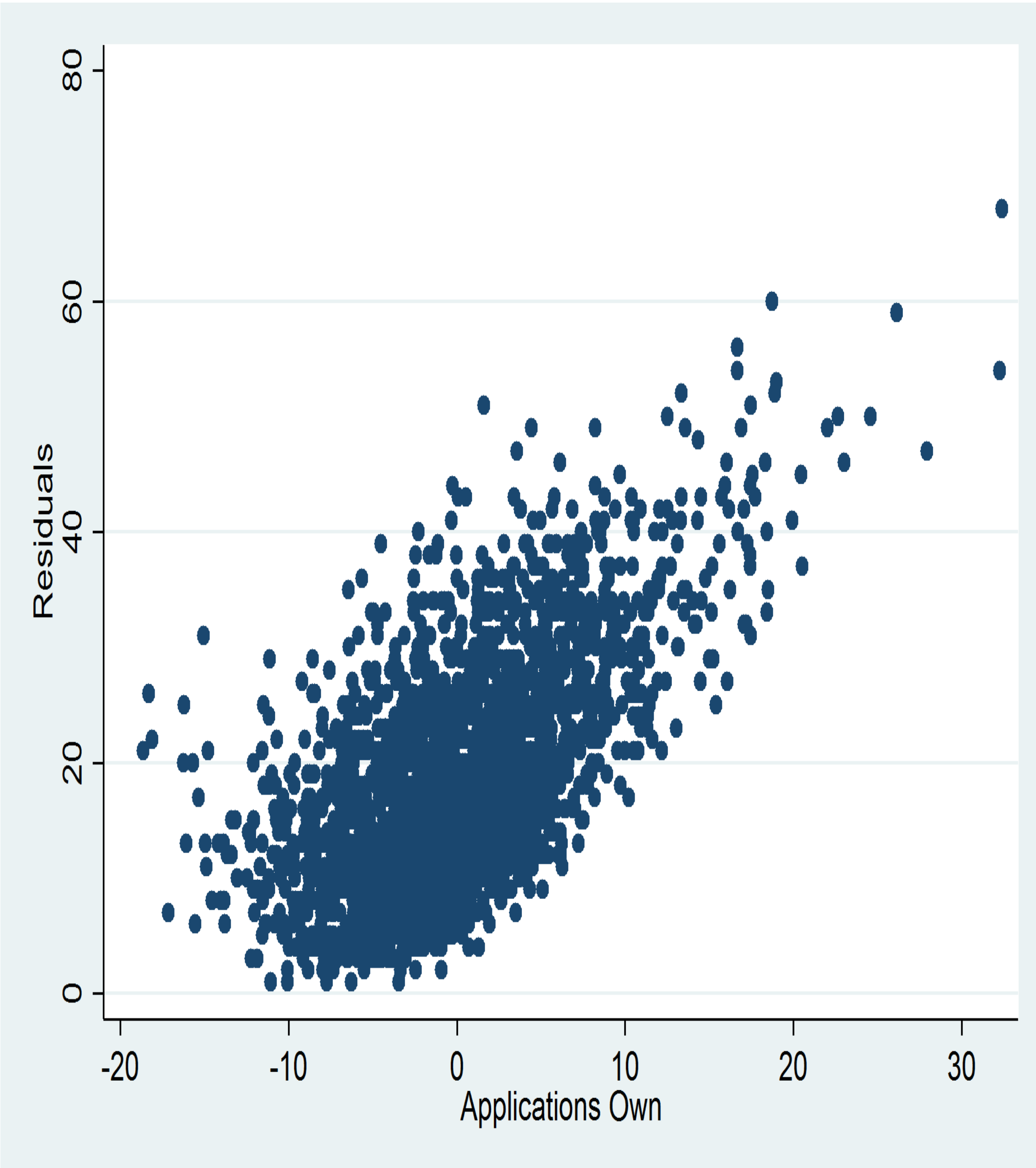


Figure 4: Residual Plot from Appendix Table