ESSAYS ON MORTGAGE SERVICING AND URBAN ECONOMICS

by

ANURAG ASHOK MEHROTRA

(Under the Direction of HENRY J. MUNNEKE and RICHARD W. MARTIN)

ABSTRACT

The first essay focuses on mortgage servicer behavior with regard to delinquent mortgage loans, and the second one concentrates on the impact of spatial mismatch of affordable housing and jobs on the labor outcomes of vulnerable groups.

In the first essay, I examine the decision of mortgage servicers to foreclose, modify, or take no action when managing delinquent loans on behalf of an investor/guarantor. It further explores if the geographic concentration of the loan holdings of the investor/guarantor influences the servicer's decision. Foreclosing on a large number of loans within a geographic area may negatively impact the value of the remaining holdings of the investor/guarantor influencing the decision to take action or not. We specifically explore the servicing of GSE loans - loans guaranteed by Fannie Mae. GSEs play an important role in the mortgage market and analyzing the role of the agency servicers is important to understand mortgage markets as a whole. The results indicate increases in geographic concentration of Fannie loans decrease the probility of foreclosure. As the foreclosure action of *non-s* Fannie servicers increases, the probability of the *non-s* Fannie servicer foreclosing on its delinquent loans increases. However, as the foreclosure action of *non-Fannie* servicers' loans increases, the probability of the *non-Fannie* servicer foreclosing on its delinquent loans decreases.

In my second essay, I examine the spatial mismatch hypothesis, which was originally formulated by John Kain (1968), suggests that there is a relationship between the relative spatial distributions of employment and population in a metropolitan area and labor market outcomes. Specifically, the hypothesis posits that labor market outcomes are negatively affected by the divergence of the two distributions. This impact is assumed to be strongest for the least mobile population groups in the metropolitan area. Spatial mismatch has proven to be remarkably persistent over time (Martin (2004)).

I pursue one possible explanation for the persistence of spatial mismatch: Jobs could be moving to areas where there are fewer affordable housing units available. If this is the case, then residents who are disadvantaged by employment shifts, will find it harder to move from areas with decreasing job availability to areas experiencing job growth.

This paper uses tract-level data on employment and affordable housing units from 2000 to 2010 for 103 U.S. metropolitan statistical areas to document the extent to which jobs have been moving to areas where there are fewer affordable housing units. The results suggest that jobs have been moving away from areas with more affordable housing units and that the shift is stronger for units that are affordable to households with the lowest income. The paper then analyzes whether there is a relationship between the degree to which jobs moved away from affordable housing units and the labor market outcomes of workers from the following four vulnerable groups – Black males, Black females, Hispanic males, and Hispanic females. The paper's results suggest that there is no evidence that spatial mismatch between affordable housing and jobs contributes to the persistence of poor labor market outcomes of Blacks and Hispanics.

INDEX WORDS: Loan Defaults, Loan Modifications, Foreclosures, Mortgages, Mortgage Servicing, Fannie Mae, Government Sponsored Enterprise (GSE).
Spatial Mismatch, Metropolitan Employment Shifts, Employment, Affordable Housing, Jobs.

ESSAYS ON MORTGAGE SERVICING AND URBAN ECONOMICS

by

ANURAG ASHOK MEHROTRA

BSc, University of Mumbai, INDIA, May 1992

PGDACM, Narsee Monjee Institute of Management Studies, INDIA, May 1995

MBA, Case Western Reserve University, May 1999

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

ESSAYS ON MORTGAGE SERVICING AND URBAN ECONOMICS

by

ANURAG ASHOK MEHROTRA

Major Professors:	Henry J. Munneke
	Richard W. Martin
Committee:	James B. Kau
	Darren K. Hayunga
	James N. Conklin
	Ruchi Singh

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2019

DEDICATION

I dedicate this dissertation to my late father, from whom I learned not only to dream big but to own the challenges that come with setting lofty goals. Knowing his vision of my earning a PhD, later sharing in that vision, and now realizing it has been the most rewarding and enriching circle I've closed!

ACKNOWLEDGEMENTS

It is not an easy undertaking to thank, in just a few words, all the individuals who made this dissertation possible. To begin with, I am indebted to my dissertation advisers, Henry Munneke and Richard Martin; their patience, support, stewardship, and continued encouragement have been invaluable during this process. I am also deeply appreciative of the insightful comments and counsel of my dissertation committee members Jim Kau, Darren Hayunga, James Conklin, and Ruchi Singh; each of them, in their own way, has provided me extensive guidance (personal and professional) at the times I've needed it the most. This journey would have been infinitely more solitary were it not for the camaraderie and reassurance of my fellow students Ryan Lowery, Jim Stevens and Nick Smith - I wish them luck in all of their endeavors!. A special thank you to my office mate Ryan Lowery; our friendship will continue long after we both graduate. I would also like to thank Robert Hoyt, who always made time to discuss my professional and personal predicaments and the staff at the Terry College of Business - Wendy Wyatt, Linda Dalton, Logistics, OITS and the student workers - for their unquestioning assistance from time to time.

The completion of this dissertation would have been impossible without the unstinting support of my family: my mother (Shobha), my brother (Abhishek), and my in-laws (Prem and Vanmala). I thank my mother for her unconditional love and her supreme confidence and eternal belief in my success. My father-in-law's optimism and hope, in turn, kept mine kindled on the toughest days. I don’t know where to begin to say thank you to my kindred soul and my best friend: my dearest wife, Promilaa Bhatia - I love you for everything. My two beautiful children, Viveka and Dheer, are my pride and joy and always inspire me to go above and beyond.

I also wish to acknowledge and thank the late Carolyn Dehring, who recruited me into the Terry College of Business. She and her family opened their doors and hearts to us at a vulnerable time in their lives, even as we embarked on a new and exciting chapter in our lives.

I couldn't be more grateful that this experience and process have prepared and presented me with the opportunity to learn from and work with the best and brightest minds in my field of study. The least I can and will do is to pay it forward.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES ix

LIST OF FIGURES xi

CHAPTER

1 LOAN SERVICING AND THE MANAGEMENT OF DELINQUENT LOANS 1

1.1 INTRODUCTION 1

1.2 THE MORTGAGE SERVICER MODEL 5

1.3 LITERATURE REVIEW 12

1.4 DATA 16

1.5 EMPIRICAL FRAMEWORK 19

1.6 RESULTS 22

1.7 ISSUES OF ENDOGENEITY 25

1.8 ESTIMATES OF THE FIXED EFFECT VARIABLES 28

1.9 CONCLUSION 30

1.10 REFERENCES 33

2 SPATIAL MISMATCH AND AFFORDABLE HOUSING 46

2.1 INTRODUCTION 46

2.2 LITERATURE REVIEW 49

2.3 SMI/ GSMI ANALYSIS 55

2.4 REGRESSION ANALYSIS..68

2.5 CONCLUSION 84

2.6 REFERENCES 86

LIST OF TABLES

Page

Table 1.1: Descriptive Statistics for the Pool of Delinquent Mortgage Loans 38

Table 1.2: The average concentration of loans by servicer across time by MSA..........................39

Table 1.3: Estimate of the Fannie Servicer's Decision with Local Loan Concentrations 40

Table 1.4: Estimate of the Fannie Servicer's Decision with Local Loan Concentrations 41

Table 1.5: Estimate of the Fannie Servicer's Decision with Local Loan Concentrations 42

Table 1.6: Estimate of the Fannie Servicer's Decision with Local Loan Concentrations 43

Table 1.7: Estimate of the Fannie Servicer's Decision with Local Loan Concentrations utilizing a Control Function (2SRI) 44

Table 1.8: Estimate of the Fannie Servicer's Decision with Local Loan Concentrations utilizing a Control Function (2SRI). 45

Table 2.1: Weighted average Spatial Mismatch Indices and the General Spatial Mismatch Indices for rental units across the 103 largest MSAs in the US..................................89

Table 2.2: Weighted average Spatial Mismatch Indices and the General Spatial Mismatch Indices for rental units across the 103 largest MSAs in the US..................................90

Table 2.3: Summary Statistics across the 103 largest MSAs in the US. 91

Table 2.4: Summary results for the impact of the general spatial mismatch index on the unemployment rate of Blacks and Hispanics. 92

Table 2.5: Summary results for the impact of the general spatial mismatch index on the employment to population ratio of Blacks and Hispanics. 93

Table 2.6: Summary results for the impact of the change in the general spatial mismatch index on the unemployment rate of Blacks and Hispanics..94

Table 2.7: Summary results for the impact of the change in the general spatial mismatch index on the employment to population ratio of Blacks and Hispanics 95

Table 2.8: Summary results for the impact of the decomposed general spatial mismatch index on the unemployment rate of Blacks and Hispanics..96

Table 2.9: Summary results for the impact of the decomposed general spatial mismatch index on the employment to population ratio of Blacks and Hispanics 97

LIST OF FIGURES

Page

Figure 1.1: Dual shock model explaining mortgage delinquency 36

Figure 1.2: Log odds of foreclosure and modification across the sample period..........................37

CHAPTER 1

LOAN SERVICING AND THE MANAGEMENT OF DELINQUENT LOANS

1.1 Introduction:

During the most recent financial crisis, and the resulting period of steep declines in house prices, the mortgage servicing industry emerged as a pivotal player in managing delinquent mortgage loans. Even before the crisis, the mortgage market had moved from portfolio lending with an 'originate-to-hold' approach to a securitized lending or an 'originate-to-distribute' model resulting in robust growth of the mortgage servicing industry.[1] This evolution in the mortgage servicing industry has garnered a lot of attention from lenders, investors, and regulators.

Under an 'originate-to-distribute' model, newly originated loans are sold in the secondary market and packaged into securities often guaranteed by Fannie Mae or Freddie Mac (government-sponsored enterprises or GSEs). However within an agency/GSE mortgage backed security (MBS) model, the credit risk associated with the loan remains with the GSE due to its guarantee of the loan. In this scenario, a principal-agent servicing arrangement between the GSEs and mortgage servicers is formed. Since the agency issuing the MBS bears all of the credit risk, the agency monitors the terms of the pooling and servicing agreements to minimize any credit loss. The current study explores the decisions of agency servicers to modify, foreclose, or take no action on delinquent loans. More importantly, it explores if agency mortgage servicers, as representative agents of investors/guarantor, internalize the potential negative price impact of foreclosure, through their servicing actions, to minimize the impact on the remaining pooled assets. With a large portion of the $10 trillion in mortgage debt being managed under

[1] Federal Housing Finance Agency: Office of Inspector General, A Brief History of The Housing Government Sponsored Enterprises, available at https://www.fhfaoig.gov/LearnMore/History

agency service arrangements, analyzing the role of the agency servicer is important to understand mortgage markets.

Favara and Giannetti (2017), within the context of portfolio lending, provide evidence that lenders who hold a large share of outstanding mortgages in a geographic area on their balance sheet tend to avoid foreclosure and seek to modify more loans. They theorize that if lenders have a large concentration of loans in a geographical area, they will foreclose less often to minimize the decline in local house prices associated with foreclosure on their remaining portfolio.[2] On the other hand, when mortgages are held by many fragmented lenders (lenders with low concentrations), each lender disregards the effects of its foreclosure decisions on local house prices resulting in higher levels of foreclosure. Favara & Giannetti conclude that portfolio lenders' propensity to foreclose on delinquent loans is negatively correlated to their share of outstanding mortgages, i.e., portfolio lenders with a high concentration have a higher propensity to modify a delinquent loan as opposed to lenders with a low concentration.

The negative spillovers associated with foreclosures are fundamental to the model theorized by Favara and Giannetti (2017) and has been examined by numerous studies. Shleifer and Vishny (1992) demonstrate that when markets are illiquid, forced sales of assets fetch prices below their fundamental values and distressed sales generate spillovers that reduce the market value of related assets held by the other players in the market. Other studies have shown that foreclosures generate price declines and have a propensity to spillover to adjoining and neighboring properties and endogenous housing illiquidity amplifies this drop in house prices (for example see Harding, Rosenblatt, and Yao, (2009); Campbell, Giglio and Pathak, (2011); Anenberg and Kung, (2013); Hartley, (2014); Garriga and Hedlund (2017)). Mian, Sufi and Trebbi (2015) showed during the recent crisis that foreclosures led to a widespread drop in house prices. This widespread decline resulted in further delinquencies as homeowners were upside down on their mortgages – they owed more than the value of their house. Guren and McQuade (2018) showed

[2] Foreclosures generate negative externalities, which could result in lower house prices in their immediate vicinity (negative spillover effect).

that during an economic downturn, lender equity is eroded, buyers are selective and the number of buyers is reduced due to declining credit scores. When combined, these factors result in lower prices and transactions in the housing market. Foreclosure sales intensify price declines and cause additional defaults, which creates further price declines and generates a feedback loop.

Although the current study evaluates the behavior of the servicers of GSE loans with regard to delinquent loans, we believe the theoretical model developed by Favara and Giannetti (2017) provides a basis for our outcomes. In the case of mortgages that are retained by lenders on their balance sheet, the credit risk is held by the lender. In the securitized market (GSE/Agency), the guarantor holds the credit risk. Hence, guarantors in the securitized market are the lenders' counterpart in the portfolio lending market. Therefore we are able to utilize the lender model for servicers. The portfolio lender will act to minimize loss, as will the guarantor. Hence, we expect Fannie servicers to act in a manner consistent with portfolio lenders and loan guarantors. The current study extends this line of research by exploring how the decisions of Fannie Mae servicers to modify, foreclose, or take no action on delinquent loans is impacted by the geographic concentration of Fannie loans they service and non-Fannie loans serviced by others . The focus of our paper is similar to Favara and Giannetti (2017) but shifts the examination from the owner-servicer (portfolio lender) to a principal-agent relationship with the servicer acting as the agent for the investor/guarantor.

To examine the mortgage servicers' behavior with respect to delinquent loans, we employ an extensive research data set released by Fannie Mae. The full data set released by Fannie Mae consists of over 22 million loans originated from 2000-2016. Monthly loan-level performance information, as well as acquisitions data are available for each loan. We restrict the data to 30-year fixed rate, fully amortizing, mortgage loans that are delinquent for more than 90 days. We also only consider loans originated in one of the cities covered by S&P CoreLogic Case-Shiller 20-City Composite Home Price NSA Index. These 20 metropolitan areas represent about one third of the 2010 U.S. population.

To mimic the servicers' behavior each month, the probability of a loan being modified or foreclosed upon is estimated using a loan hazard model based on monthly loan performance data. Specifically, we employ a discrete-time hazard model, estimated using a multinomial logit model, which addresses issues of left truncation and right censoring. Upon delinquency, each month the servicer has the option to either foreclose on the delinquent mortgage and liquidate the asset, or modify/renegotiate the delinquent mortgage such that the borrower resumes the repayment of the mortgage, or take no action on the delinquency pending a decision. Given the three primary outcomes chosen by a servicer, each delinquent observation continues until either an event (foreclosure or modification) occurs or the observation is censored. It is important to note that a servicer's revenue stream is a function of the length of the mortgage (duration). When a mortgage becomes delinquent, choosing to foreclose on the mortgage will suspend servicing revenue while modifying it will preserve the revenue stream. During the delinquency period, servicers are contractually obligated to provide short term financing (the servicer has to advance the monthly principal and interest payments) to a delinquent mortgagor until the delinquency is resolved (execution of a modification or the sale of the foreclosed house). Upon resolution, the servicer is reimbursed all previously made advances on that delinquent mortgage loan irrespective of the terms of the modification or foreclosure proceeds.

The results indicate that as the dollar amount of mortgages serviced by Fannie servicers in a Metropolitan Statistical Area (MSA) increases, the probability that the delinquent loan will be foreclosed decreases. We also find that as the foreclosure action of other Fannie servicers in an MSA increases, the probability a delinquent loan will be foreclosed by a Fannie servicer increases. In other words, Fannie servicers follow the action of the other Fannie servicers and disregard the effects of its foreclosure decisions on local house prices resulting in higher levels of foreclosure. Interestingly, as the foreclosures of non-Fannie loans increase, the probability of no action increases and that of modification and foreclosure decrease. The foreclosure actions of non-Fannie servicers doesn't appear to impact the foreclosure action of Fannie servicers.

As with most studies, there are potential issues related to endogeneity within this study, most directly, is Fannie Mae's choice to aggressively pursue certain markets, which directly impacted the dollar magnitude of Fannie loans outstanding in a particular geographic area. The pursuit and exposure in certain markets by Fannie Mae is a direct consequence of their management of political risk in conjunction with the actions of the House Appropriations Committee. Hence, we use the tenure of House representatives on the House Appropriations Committee from the MSA as an instrument to predict the total number of Fannie loans serviced in an MSA. Nevertheless, we employ a control function estimation method, as suggested by Terza, Basu & Rathouz (2008), in our final model to minimize any potential endogeneity bias.

The rest of the paper is organized as follows. In section 1.2, we adjust Favara and Gianetti's portfolio lender model to reflect a mortgage servicer model and in section 1.3 compile a comprehensive literature review and adjust Favara and Gianetti's portfolio lender model to reflect a mortgage servicer model. Section 1.4 describes the data and section 1.5 the empirical framework. Sections 1.6 summarizes the results and section 1.8 concludes.

1.2 The mortgage servicer model:

Favara and Gianetti (2017) develop a cogent model that explains how foreclosures, house prices and the concentration of mortgages are related. The underlying assumption in their model is that foreclosures result in neighboring house price declines. In their model, an income shock causes a disequilibrium between supply and demand of housing. Prior to the income shock, the housing demand for a household is determined such that the utility from owning a house is greater than the price of housing. The equilibrium housing price is determined when the aggregate supply of housing is equal to aggregate demand of housing. Under this equilibrium all households repay their mortgage debt and there are no delinquencies.

When a segment of households are impacted by an economic shock, the affected households cannot repay their mortgage debt. When lenders foreclose on these delinquent mortgages, these households are excluded from the owned housing market and become renters and their foreclosed houses will be on the market. This results in an imbalance and equilibrium house price after the shock is strictly lower than prior to the shock and decreases in the probability of foreclosure i.e. the greater the probability of foreclosure the lower the price of houses in equilibrium.

The lender chooses to maximize their payoff (choose the option that maximizes the net present value) from their portfolio of performing and delinquent mortgages. Consequently, when lenders hold a large share of mortgages they are more likely to modify delinquent mortgages in an effort to internalize the negative externalities caused by foreclosures, whereas, lenders that hold a small share of mortgages disregard the negative externalities caused by foreclosures and always foreclose delinquent mortgages. In the case of multiple large lenders, each lender's probability of foreclosure increases in the share of the mortgages retained by the other large lenders. In conclusion, foreclosures are inversely related to the overall concentration of outstanding mortgages on lenders' balance sheets.

While borrowing heavily from the Favara and Gianetti (2017) model, we depart from it on two counts. First, an income shock is not a sufficient condition for default, rather negative equity is necessary. Hence, we introduce delinquencies in our model by employing a dual shock model – housing price shock and an income shock. Even though we arrive at the same outcome, the path to delinquency is different. Second, we evaluate the behavior of mortgage servicers as opposed to portfolio lenders. Mortgage servicers as "agents" are conflicted between maximizing the expected payoff for their "principal" (fiduciary duty) and the revenue from mortgage servicing (Levitin and Twomey, 2011). This conflict may impact the servicer behavior.

Mortgage delinquency can be explained using the 'equity' approach or the 'ability to pay' approach. Under the 'equity' approach, the decision to become delinquent (and eventually default) is seen as the household (mortgagor) exercising their put option, i.e., put the mortgage back to the lender at the outstanding balance of the mortgage. The household exercises this put option when the cost of default is lower than the household's negative equity in the property (the put option is in the money). The 'ability to pay' approach focuses on the events that 'trigger' delinquency. For example, 'trigger' events could include loss of wages, divorce, ill-health, etc. It is important to note here that that negative equity is a necessary, but not sufficient condition for delinquency (eventual default). Hence, in our model we assume dual shocks causing mortgage delinquencies – a price shock and an income shock.

Scenario 1: Negative Housing Price Shock (Single Shock)

At t=1, there are two states of the world. In the bad state of the world, a fraction e of households suffer a negative housing price shock and their outstanding mortgage is greater than the value of their house. In the good state of the world, the remaining fraction $(1 - e)$ of households' outstanding mortgage is equal to or lower than the value of their house. We assume that price shocks are observable, and that the households whose outstanding mortgage is greater than the value of their house at delinquency, i.e., $M > p^D$, exercise the put option on the house (which is in the money) and stop paying their mortgage. Households with an outstanding mortgage equal to or lower than the value of their house continue repaying M. At this point, the servicer forecloses on a fraction λ of delinquent mortgages and sells the houses at price p^F and modifies/renegotiates the remaining fraction $(1 - \lambda)$ of delinquent mortgages such that $p^F < M' < M$ to induce the borrower to continue repaying M'.

Scenario 2: Negative Housing Price Shock coupled with a Negative Income Shock (Dual Shock)

At t=1, there are two states of the world. In the bad state of the world, a fraction e of households suffer a negative price shock and/or income shock (receive zero or reduced wages), while the remaining fraction continues to receive w. In the good state of the world, all households receive w. We assume that both price and income shocks are observable, and only part of the households' income can be pledged to make the mortgage payment and the households whose outstanding mortgage is greater than the value of their house and the income shock triggers a delinquency, i.e., $M > p^D$.

Under 'the dual shock scenario' assumptions, households that have an outstanding mortgage greater than the value of their house $M > p^D$ and suffer an income shock $w = 0$, stop paying their mortgage. At this point, the servicer forecloses on a fraction λ of delinquent mortgages and sells the houses at price p^F (figure 1 – outcome 1), modifies/ renegotiates the remaining fraction $(1 - \lambda)$ of delinquent mortgages such that $p^D < M' < M$ to induce the borrower to continue repaying M' (figure 1 – outcome 2) or the servicer holds the mortgage in abeyance pending a decision (figure 1 – outcome 3). For households that suffer an income shock ($w = 0$) but have an outstanding mortgage equal to or lower than the value of their house $M \leq p^D$ can pay off the mortgage by selling their house and the foreclosure is theoretically avoided (figure 1 – outcome 4). Households with high income realization (continue to receive w) and an outstanding mortgage equal to or lower than the value of their house $M \leq p^D$ continue repaying their mortgage (figure 1 – outcome 5). Under all scenarios, the necessary but not sufficient condition for a delinquency/default is that the outstanding mortgage balance is greater than the value of the house i.e. $M > p^D$.

Next, like Favara and Giannetti (2017), we determine the equilibrium house price and servicers' foreclosure probabilities upon delinquency. The shift from portfolio lending to securitized lending has delegated the responsibility to manage mortgages to the servicer in return for a servicing fee. When a mortgage becomes delinquent, the servicer's choice to foreclosure or modify a mortgage must be driven by their fiduciary duty to the investor by choosing the option that results in a higher net present value. Foreclosures create a negative externality, which results in lower house prices. By foreclosing fewer mortgages, the servicer will minimize the negative externality and protect the investor. Since the servicer does not own the mortgage and is paid a fee to manage the loan on behalf of the investor, the servicer may be conflicted between maximizing the value of the mortgage and their revenue stream.

On the surface, mortgage servicers, as opposed to lenders/investors, do not fund the mortgages and hence do not face the risk of repayment/loss as a result of foreclosure. In other words, servicers have little incentive to modify delinquent loans. However, a servicer's revenue stream is a function of the length of the mortgage (duration). Moreover, when a mortgage becomes delinquent, choosing to foreclose on the mortgage will suspend servicing revenue while modifying it will preserve the revenue stream. During the delinquency period, servicers are contractually obligated to advance the monthly principal and interest payments (the servicer provides short term financing to the delinquent mortgagor) to the investor until the delinquency is resolved (execution of a modification or the sale of the foreclosed house). Upon resolution, the servicer is reimbursed all previously made advances on that delinquent mortgage loan irrespective of the terms of the modification or foreclosure proceeds.

The servicer maximizes their own servicing revenue with a combination of revenue from performing loans, modified loans and foreclosed loans represented by equation 1 below:

$$\max_{\lambda}\{\phi_1(1-e)M + \phi_2 e(1-\lambda)M' + \phi_3 e\lambda p^F\} \quad (1)$$

where ϕ_k is the fee from servicing the three different mortgage types (performing, modified and foreclosed). The first term represents the servicing fee from performing loans which is a function of the mortgage balance. The second represents the servicing fee from performing modified loans, which is a function of the modified mortgage balance. The third represents the servicing fee from foreclosing defaulted loans which is a function of the house value. The servicer picks λ – the percentage of delinquent loans that will be foreclosed – such that their own revenue is maximized.

As an agent for the investor/guarantor, the servicer should select the λ that maximizes the value of the investor's portfolio's value. If the servicer disregards the impact of concentration in a market the servicer's decision, λ will be closer to 1 and if the concentration in a market affects servicer's decision, λ will be closer to 0. Therefore, if a servicer's decision on delinquent loans isn't impacted by concentration, the servicer forecloses with a probability closer to 1 to maximize their revenue, their foreclosure actions will result in declining house prices, which will result in a decline in the net present value of the investors' portfolio or an increase in loan losses to the guarantor. Alternatively, if the concentration in a market affects the servicer's decision, fewer foreclosures result in stabilizing house prices and maximizing the value of the investor's portfolio.

Extending the model to include one large servicer and many small servicers, smaller servicers continue to foreclose with probability one to maximize their payoff. On the other hand, a large lender internalizes their foreclosure decision to maximize their payoff, i.e., the large lender's foreclosure probability is less than one. In this case, the equilibrium house price is, ceteris paribus, higher than in the case of atomistic lenders, allowing the lenders who hold a large share of mortgages to maximize their payoff on delinquent mortgages by foreclosing fewer delinquent mortgages. This results in fewer foreclosures and correspondingly higher house prices following negative shocks than in markets with just smaller servicers.

The large servicer maximizes the same revenue maximization function of a smaller servicer, but internalizes the effect of its foreclosure policy on housing prices. Thus maximizing $\max_{\lambda}\{\phi_1(1-e)M + \phi_2 e(1-\lambda)M' + \phi_3 e\lambda p^{F'}\}$ is equivalent to solving the following problem:

$$\max_{\lambda_1}\{\phi_1(1-e)M + \phi_2 e(1-\lambda_1)M' + \phi_3 e\lambda_1\left(\bar{y} - \frac{\bar{H}}{1-e} + \frac{(1-\lambda_1)\zeta_1 e\bar{H}}{1-e}\right)\} \tag{2}$$

The optimal λ_1, fraction of delinquent loans the large servicer forecloses is:

$$\lambda_1 = min\left\{\frac{1}{2}\left(1 + \frac{\frac{(1-e)(M'+\bar{y})}{\bar{H}} - 1)}{e\zeta_1}\right), 1\right\} \tag{3}$$

Suggesting the concentrated servicer's optimal foreclosure probability $\lambda_1 < 1$ when the share of mortgages held by the servicer, ζ_1, is sufficiently large. $\lambda_1 < 1$ when the fraction of households affected by the economic shock, e, is large, housing demand relative to the size of the housing stock, $\bar{y}/\bar{H}$, is large and the number of modified mortgages relative to the size of the housing stock, $M'/\bar{H}$, is large.

Hence, mortgage servicers with large concentrations in a geographical area are less likely to foreclose on delinquent loans resulting in lesser foreclosures in that area. Thus, housing prices will be higher represented by $p^{F'}$ as opposed to p^F and the value of the investor's portfolio will be maximized. Hence, we must observe servicers that have a high concentration of loans in a geographical area to display a lower probability to foreclose a delinquent mortgage and a higher probability to modify a delinquent mortgage.

In conclusion, lenders with a large share of outstanding mortgages have an incentive to internalize the effects of foreclosures on house prices. Thus, we have the following hypotheses:

Hypothesis 1a: When the dollar amount of mortgages serviced by a servicer in an MSA increases, the probability that the delinquent loan will be modified increases and the probability that the delinquent loan will be foreclosed decreases. As the concentration of the dollar amount of mortgages managed in an MSA by a servicer increases, the servicer will foreclose with a probability less than one on a delinquent loan to minimize the decline in local house prices associated with foreclosures on their remaining portfolio.

Hypothesis 1b: Large servicers will foreclose with a probability less than one on a delinquent loan to minimize the decline in local house prices associated with foreclosures in their remaining portfolio. Small servicers will foreclose with a probability equal to one on a delinquent loan and will be indifferent to the decline in local house prices associated with foreclosures on their remaining portfolio.

When we include multiple large servicers and many small servicers, small servicers continue to foreclose with probability one to maximize their payoff. In this scenario, the foreclosure probability is greater than in the case of one large servicer. The equilibrium house price continues to be inversely related to foreclosure probability. If the share of the large servicers is constant and the number of large servicers increases, the concentration of each servicer reduces, which results in the probability of foreclosure increasing. Alternatively, if the number of large servicers decrease, the concentration of each servicer increases and the foreclosure probability correspondingly decreases.

The N "large" servicers maximize the same revenue maximization function of a servicer, but internalize the effect of their foreclosure policy on housing prices.

$$\max_{\lambda_1}\{\phi_1(1-e)\textstyle\sum_{i=1}^{N} B + \phi_2 e(1-\lambda_1)\sum_{i=1}^{N} B' + \phi_3 e\lambda_1 p^{F''}\} \tag{4}$$

$$\max_{\lambda_1}\{\phi_1(1-e)\textstyle\sum_{i=1}^{N} B + \phi_2 e(1-\lambda_1)\sum_{i=1}^{N} B' + \phi_3 e\lambda_1\left(\bar{\gamma} - \frac{\bar{H}}{1-e} + \frac{(\zeta - \sum_{i=1}^{N}\zeta_i\lambda_i)e\bar{H}}{1-e}\right)\} \tag{5}$$

$$\lambda_i = min\left\{\frac{N}{(N+1)}\left(1 + \frac{\frac{(1-e)(\sum_{i=1}^{N} B' + \bar{\gamma})}{\bar{H}} - 1)}{e\sum_{i=1}^{N}\zeta_i}\right), 1\right\} \tag{6}$$

In the scenario of N large servicers, λ continues to be inversely related to the individual servicer's share ζ_i of outstanding mortgages. If the other large servicer picks a $\lambda < 1$, and modifies a delinquent loan as opposed to foreclosing it, ceteris paribus the equilibrium house price will be larger. This will make it optimal for the large servicer to foreclose on a delinquent loan to maximize the payoff for their investor. Thus the modification decision of the other servicers results in a free rider problem, which is eliminated by the other large servicer choosing to foreclose rather than modify the delinquent loans. Therefore, each large servicer's probability of foreclosures increases in the share of the mortgages foreclosed by the other large lenders.

Hypothesis 2: When the dollar amount of mortgages foreclosed by other servicers in an MSA increases, the probability that the delinquent loan will be foreclosed increases because, by foreclosing, the large servicer is maximizing the payoff to the investor. Each large servicer takes into account the foreclosure action of the other large servicers and disregards the effect of its foreclosure decision on local house prices resulting in higher level of foreclosures.

1.3 Literature review:

The link between foreclosure and neighborhood price effects is an underlying condition of the current study. An extensive literature exists supporting the idea that foreclosures generate negative price effects on neighboring properties. Homeowners with delinquent loans may lack the resources necessary to maintain their property leading to the deterioration of the property. Foreclosure may lead to vacant or abandoned properties, which may result in physical blight within a neighborhood. Abandoned homes could provide a safe haven for criminal activities and vandalism which may negatively impact neighborhood quality. Finally, an increase in foreclosures increases the inventory of available houses in the short term and given stable demand for housing, prices of all houses in that price point will be negatively affected.

Immergluck and Smith (2006) measured the impact of the proximity of foreclosure activity (in 1997 and 1998) on 9,600 single-family dwellings in Chicago sold in 1999. They concluded that a foreclosure on average results in a decline in value of 0.9% (houses within an eighth of a mile of a foreclosure). They use a hedonic regression to estimate the impact of foreclosures on the value of nearby single-family properties while controlling for characteristics of properties and their respective neighborhoods; their paper finds that foreclosures have a significant effect on the values of nearby properties.

Lin, Rosenblatt and Yao (2009) examine the negative spillover effect of a foreclosure on neighborhood property values over the housing cycle, the time period (length) of foreclosure and the distance from the subject property. Using data on foreclosed properties between 1990 and 2006 and 2006 non-distressed sales in the city of Chicago, they found the spillover effect of a foreclosure is a decline in value of 8.7% when the foreclosure is within 300 feet and 2 years of the sale. The negative spillover effect diminishes over time and distance. Similarly, Harding, Rosenblatt and Yao (2009), using a repeat model, estimate the spillover effects across seven metropolitan statistical areas and find that the greatest effect from the closest foreclosures is about 1% over and above the local trend in house prices.

Campbell et al (2011) show that houses sold after foreclosure are sold at lower prices (on average 27% lower) than other houses for sale (death related or bankruptcy related). Using a difference-in-difference approach, they show that foreclosures appear to have negative externality on the values of neighboring non-distressed properties, worsening the decline in house prices. Their study finds that a foreclosure reduces the value of houses by approximately 1 percent within a small radius of the foreclosure. They used a data set of all residential transactions from the state of Massachusetts during the period from 1987 to 2009.

Anenberg and Kung (2013) evaluate the propensity for home sellers to adjust their list price surrounding a new, nearby foreclosure listing in the San Francisco, Washington D.C., Chicago, and Phoenix MSAs over the period 2007–2009. They find that sellers reduce their listing prices in the week that real estate owned (hereinafter referred to as REO) properties are listed, which is attributed to the seller lowering their asking price due to the entry of a new REO listing.

Further, macro-economists have studied the impact of foreclosures during the recent housing bust and have explained how the resulting drop in house prices is transmitted through the economy resulting in the length and breadth of the recession. Mian, Sufi and Trebbi (2015) examine the effect of foreclosures on the prices of houses using variation in state foreclosure laws as an instrument. They find that foreclosures result in house price declines and diminish housing wealth, which in turn reduces the consumption of durable goods and impedes residential investment. They conclude that foreclosures played a significant role in the severity and duration of the housing downturn.

Garriga and Hedlund (2017) delineate the determinants of the housing crash between 2006 and 2011 using a quantitative model. They find that high risk of losing household income and tightening credit availability were the main drivers of the crash. The decline in house prices and the presence of mortgage debt increased the time on the market (housing illiquidity) which resulted in high risk of foreclosure affecting credit availability and access.

Guren and McQuade (2018), in the backdrop of the recent housing downturn, present a macro economic model that depicts how foreclosures intensify the drop in house prices and defer the recovery of the housing market. In an economic downturn, lender equity is eroded, buyers are selective and the number of buyers is reduced due to declining credit scores resulting in lower prices and transactions in the housing market. Foreclosure sales intensify price declines and cause additional defaults, which creates further price declines and generates a feedback loop. They find that a buyer's blemished credit record and the buyer's heightened pickiness account for 22.5% of the decline in house prices. Another 30% of the decline is house prices is attributed to the reduced lending as a result of the default losses suffered by the lender.

A large number of papers have evaluated and explained the differences in the modification/loan renegotiation rates and foreclosure rates between portfolio lenders (bank loans) and securitized loans.

Past research has presented three main theories that might explain the low rate of modifying delinquent mortgage loans (securitized) (i) “agency theory” (Cordell et al (2008) and Levitin and Twomey (2011)), (ii) “information asymmetry” (Foote et al (2009) and Adelino, Gerardi and Willen (2013)), or (iii) “institutional factors” (Piskorski et al (2010), Agrawal et al (2011), Been et al (2013) and Kruger (2016)). There isn’t complete unanimity among researchers as to the underlying causes of the low rate of loan renegotiation among delinquent securitized loans.

Piskorski, Seru and Vig (2010) found a significantly lower foreclosure rate among bank held delinquent loans as opposed to mortgages that were securitized. Controlling for contract terms and regional conditions, they find that seriously delinquent loans that are held by the bank (portfolio loans) have lower foreclosure rates than comparable securitized loans (between 3% (13%) and 7% (32%) in absolute (relative) terms). Subsequent research by Agrawal et al (2011), Been et al (2013) and Kruger (2016) obtain results similar to Piskorski et al (2010) and provide further evidence that bank held delinquent loans were much more likely to be renegotiated over securitized delinquent loans. These papers attribute this behavior to frictions in the securitization market that prevent servicers from modifying loans.

However, Adelino, Gerardi and Willen (2013) find that there is no difference in the rates of loan modification between securitized loans and loans held in a portfolio. At the start of the crisis, private label securities loans had a higher chance to be modified while government sponsored enterprise loans had a lower chance to be modified, and as the crisis progressed, this probability reversed positon. They offer an alternative theory that argues that due to uncertainty with regard to ability of a borrower to repay and redefault probability of a loan that has been modified, foreclosure will be chosen by the lender/servicer over renegotiation even when the cost of foreclosure exceeds the direct cost of modifying the loan to the lender/investor.

Foote et al (2009) observe that modifying delinquent loans might be more expensive to investors than foreclosing on them. Foote et al (2009) conclude that even though lenders may lose money by foreclosing on a delinquent loan, but renegotiated loans will have negative or lower NPV over a foreclosure especially if the modification is offered to borrowers who were likely to have the ability to self-cure the delinquency.

Cordell et al (2008) present evidence that servicers do not have the right incentives to modify loans. They find that modifying loans is costly for servicers because they lack adequate staff and technology and have little to no financial incentive to bolster their capability to mitigate losses for investors. Further, the investors in these private-label mortgage backed securities were concerned that the redefault rate of delinquent subprime borrowers might be high resulting in repayment delays and increased costs.

Levitin and Twomey (2011) outline a number of reasons why servicers prefer foreclosure over modification and their actions might not always be in the best interest of the investors. First, servicing contracts allow servicers to collect payment advances and fees from the proceeds of the foreclosure sale before the investors are paid. Second, servicing delinquent mortgages is more expensive than servicing performing mortgages. Hence, to maximize their revenue, servicers prefer foreclosure to modification of a delinquent mortgage.

1.4 Data:

The data for this study are drawn from Fannie Mae's single-family historical loan performance database. The data are available in two distinct sets of files: (i) 'Acquisition' files include static mortgage loan data at the time of the mortgage loan's origination and delivery to Fannie Mae, and (ii) "Performance' files provide monthly performance data for each loan, from acquisition up until its current status as of the previous quarter. When restricted to 30-year fixed rate, fully amortizing, mortgage loans, the sample consists of over 22 million loan origination records with monthly loan-level performance data beginning in Q1 2000 through Q1 2016[3]. Of these 22 million records, nearly 1.2 million became severely delinquent (90+ days past due) during the reporting period.

[3] For each loan, Fannie Mae reports the terms of originating the loan, which include loan amount, the contract rate, the origination date, and the originating lender. In addition, the following characteristics of the loan and borrower at origination are

To construct our data set, we restrict the sample to severely delinquent loans (90 days+) located in one of the 20 MSAs covered by the S&P CoreLogic Case-Shiller Home Price Indices. This focus allows us to incorporate house price changes, an important element, into our study. We also restrict our sample to servicers servicing more than one percent of the current actual total unpaid principal balance for the last month of a given quarter. Above this threshold, the Fannie Mae data identify the servicer by name, however, servicers falling under this threshold are recorded by the generic term "Other"[4]. We also remove data prior to Q1 2002 to reflect the fact that the servicer name was often left blank due to lack of servicer information during this period. Removing these loans results in the severely delinquent loan data set containing 359,388 loans of which 148,438 received no action from their servicers (they continue to stay delinquent), 67,651 were modified, and 143,129 were foreclosed. Table 1.1 provides the descriptive statistics of the severely delinquent loan data set, as well as for the sub-samples of foreclosed, modified, and no action (uncured). While the descriptive statistics in Table 1.1 are based on originated loans, it should be noted that when combined with the monthly performance data, we have over 5,613,430 monthly loan observations.In addition to Fannie Mae, we obtained the Market Trends database from CoreLogic for the period of our studay. This data set includes information on the loans outstanding and the number of foreclosures each month across the 20 MSAs. This data was crucial in enabling us to estimate the impact of all loans in the MSA on the decision of a Fannie servicer to modify or foreclose a delinquent loan.

There are some slight differences across the subsamples of severely delinquent loans, but they also share many similar traits. The average loan amount of loans at origination is $211,225.78 with loan to value ratios at origination averaging 77.24%. Interestingly, loans that were modified had the largest loan

also provided: loan to value, debt to income ratio, number of borrowers, borrower credit score, location (MSA and 3-digit zip), loan purpose, and if the buyer is a first time homebuyer. After the loan is originated, Fannie Mae provides monthly performance data on each of these loans which includes the name of the servicer, the outstanding monthly balance, loan delinquency status, date the loan's balance was reduced to zero, reason the loan's balance was reduced to zero, date the loan was foreclosed upon, date the REO property was disposed, net sales proceeds from the disposition, holding costs of REO property, and the principal amount written off.

[4] Changes that occur in the loans servicing are reflected in the data.

at origination. Foreclosed loans had a higher loan LTV (just over 80%) at origination than loans that were modified and no action was taken (~75%). The debt-to-income ratio averaged around 0.40 for all loans at origination. For loans with private mortgage insurance, the average level of coverage is 24%.

The reduction in borrowers' equity from origination to the loan becoming severely delinquent is reflected in the average 12.54 percentage point increase in the LTV. To estimate the current loan to value ratio (CLTV) for each property, the property value origination was adjusted by the appropriate S&P CoreLogic Case-Shiller Home Price Index and divided into the current outstanding balance of the loan. As expected, the increase in the LTV was greatest for loans that were ultimately foreclosed (80% to 100%) and least for loans left uncured (75% to 79%). The increase in the LTV was driven by the decline in house prices. On average, the loan balances in each subsample declined by 6.0% (approx.) from origination to delinquency and the loans were roughly the same loan age (time since delinquency).

The property types underlying the loan may impact the action taken by the servicer. A vast majority, over 70%, of the loans in the sample are for single family home (1-4 unit) with manufactured housing and co-operative loans representing a very small fraction of the loans. The percentage of loans for condominiums and planned urban developments (PUD) are 12% and 15% of the full sample, respectively.

Based on the reported averages, it appears that servicers were less likely to foreclose on homes used as a primary residence and more likely to foreclose on investor properties. Primary residences represent the highest percentage of properties underlying modified loans, with foreclosed loans having the lowest percentage of loans on primary residences. Loans for investment properties represent 7% of the loans in the delinquency sample. The share of loans on investment properties in the foreclosure subsample is over 10%.

The purpose of the loan at origination is divided between a purchase and a refinance – 31.9% of the delinquent loans were made to borrowers to purchase a property. The balance, 68.1% loans, was made to borrowers who were refinancing their mortgage loan. Of the loans for refinancing, two-thirds were part of a cash out refinancing, while the remaining one-third were not cash out refinances. The channel by which loans are originated may reflect unobservable traits of the borrower or the underlying traits of the loan itself. Of the loans in the sample, over 45% are originated by a party other than a mortgage loan seller (correspondent) and then are sold to a mortgage loan seller. Only 31.76% of the mortgages are originated by the mortgage loan seller who takes the mortgage loan application and then processes, underwrites, funds, and delivers the mortgage loan to Fannie Mae. The remaining loans (22.97%) are originated through a mortgage broker.

1.5 Empirical Framework:

Geographic Concentration of Servicing Activity - To compute the monthly concentration/exposure of Fannie servicers in a given geographical area, we need to know the Unpaid Principal Balance (UPB) of the loan each month. Fannie Mae provides the loan amount at origination rounded to the nearest thousand. It also contains the monthly unpaid principal balance of the loan six months after acquiring the mortgage loan, as well as for the remaining loan term.[5] To populate the UPB for these missing periods (the first 6 months), we estimate the monthly payment based on the first recorded monthly UPB, contract rate, and the remaining months to maturity. We then find the PV of payments over the term of the loan. If PV of the payments is within +/- $500 of the loan amount at origination we use the calculated payment to determine the unpaid balance for the first 6 months of the loan. If the balance does not fall within $500, we calculate a series of payments by adding and subtracting .05 (up to 0.5 months) to the remaining months to maturity term. Next we estimate the original balance for each of the payment. If more than one of the balances falls within $500 of the original balance, one payment was randomly selected to calculate the missing mortgage balances. We used this approach to populate the UPB for the first six months. Thereafter, we use the UPBs for each month, as provided by Fannie Mae in their database. Once we have the UPBs and the servicer's name, we are able to calculate the dollar magnitude of the loans in the

[5] In addition, for a mortgage loan that has been liquidated within six months of acquisition, the current actual UPB displays the original UPB as a rounded value.

servicer's portfolio across each servicer and for various geographic levels (e.g., national, MSA, three digit zip codes etc.).

Using this information, we are able to find the geographic concentration of the servicer's portfolio as follows:

$$Loan\_Conc_{s,m,t} = \frac{Loans\ Serviced_{s,m,t}}{Loans\ Serviced_{m,t}} \quad (7)$$

where $Loans\ Serviced_{s,m,t}$ represents the dollar amount of all Fannie loans serviced by servicer *s* in geographic area *m* during the time period *t*. $Loans\ Serviced_{m,t}$ represents the dollar amount of all loans serviced by all servicers in geographic area *m* during the time period *t*.

We are also able to find the geographic concentration of Fannie loans as follows:

$$Loan\_Conc_{f,m,t} = \frac{Loans\ Serviced_{f,m,t}}{Loans\ Serviced_{m,t}} \quad (8)$$

where $Loans\ Serviced_{f,m,t}$ represents the dollar amount of all Fannie loans serviced in a geographic area *m* during the time period *t*.

The average concentration of loans within the 20 MSA is surprisingly similar. Table 1.2 contains descriptive statistics of the concentration of loans, as defined in equation (7 and 8), per MSA. The averages in the table are found by calculating each servicer's concentration in an MSA at each point in time, averaging these concentrations across servicers at each point in time, and then finding the average over all time periods. The average servicer's holdings are 7% of the Fannie Mae serviced loans in an MSA with a range being from 4.71% to 7.74%. Combining this with the median number of servicers in an MSA of 11.5, these top servicers service roughly 80% of all the Fannie Mae serviced loans in an MSA. However, when the study period is divided into sub-periods, the data reveal the average percentage of loans held by a servicer declines substantially relative to the first five years of the study to the end.

Empirical Framework - To model a Fannie servicer's monthly decision related to the delinquency of a loan, a Cox discrete-time competing-risks loan hazard model is used. For each month of each loan subsequent to loan delinquency, loan performance data are used to determine whether a loan is continued, modified, or foreclosed. This discrete-time model solves the issue of left truncation and right censoring which are common issues in the mortgage literature. We implement the model by estimating a multinomial logit model using monthly mortgage observations of delinquent loans.[6] This framework mimics the servicer's monthly decision to modify, foreclose, or allow the delinquent loan to continue. Such a model can be written as:

$$\Psi_{i,m,t} = \alpha_0 + \beta_1 Loan\_Conc_{f,m,t} + \beta_2 \boldsymbol{X}_{i,m,t} + \delta_s + \gamma_m + \tau_t + \theta_t + \epsilon_{i,s,m,t} \tag{9}$$

where $\Psi_{i,s,m,t}$ equals 1 if delinquent loan i held by servicer s in geographic region m at time t is modified, 2 if delinquent loan i held by servicer s in geographic region m at time t is foreclosed upon, and 0 otherwise (i.e. no action is taken by the servicer). The vector $\boldsymbol{X}_{i,m,t}$, includes loan–level and borrower characteristics, some are time varying (e.g., current LTV, months since delinquency) and others are observed at the time of mortgage origination (e.g., origination channel, property type). Many of the variables are presented in Table 1.1. We also include variables representing if the borrower is a first time homebuyer, if the loan seller is equal to the loan servicer, the spread between the 10yr US Treasury and the 6 month lagged 10yr US Treasury, and the spread between the current 10yr US Treasury and 10yr US Treasury at origination. Several sets of fixed effects are also utilized including servicer fixed effects (δ_s), geographic location fixed effects (γ_m), origination year fixed effects (τ_t), and year of delinquency fixed effects (θ_t). Note when MSA fixed effects are used, they also proxy for state fixed effects as most MSAs are in a single state except for 3 in California and 2 in Florida.

While it is not uncommon to model a concentration as a ratio, empirically it might be more informative to allow the components of the concentration variable to have separate effects. In other words, to disentangle the impacts of each variable, we replace the concentration ratio with its two parts -

[6] The multinomial logit model embraces the independence of irrelevant alternatives (IIA) assumption that the odds ratio for any pair of choices is assumed to be independent of any third alternative (one event is not informative to the other conditional on all of the covariates in the model), and choices at any point in time are independent of those at any other point in time. We test this assumption and find it is an appropriate assumption for our model.

the dollar amount of mortgages serviced by a servicer *s* in geographic region *m* at time *t* and the dollar amount of loans serviced by all servicers in geographic region *m* at time *t*. Such a model can be written:

$$\Psi_{i,s,m,t} = \alpha_0 + \beta_1^n Loans_{Serviced\,f,m,t} + \beta_1^d Loans_{Serviced\,m,t} + \beta_2 \boldsymbol{X}_{i,m,t} + \delta_s + \gamma_m + \tau_t + \theta_t + \epsilon_{i,s,m,t} \tag{10}$$

where β_1^n and β_1^d represent the separate effects of the concentration of loans.

1.6 Results:

To examine the influence of loan concentration on the Fannie servicer's decision to modify, foreclose, or leave a delinquent loan uncured, we estimate a MNL model, eq. (10), including concentration variables representing the dollar amount of mortgages serviced by a Fannie servicers in MSA *m* and the dollar amount of mortgages serviced by all servicers in MSA *m* as measures of concentration. The concentration at the MSA level is the most localized geographic area in the current study. The results, reported in Table 1.3, indicate that an increase in the dollar amount of mortgages serviced by Fannie servicers in an MSA decreases both the relative probability that the delinquent loan will be modified vs. no action and the relative probability that the delinquent loan will be foreclosed vs. no action. On the other hand, an increase in the dollar amount of delinquent loans serviced by All servicers in MSA *m* increases the relative probability that the delinquent loan will be modified vs. no action and decreases the relative probability that the delinquent loan will be foreclosed vs. no action. Table 1.3 also contains the estimated marginal effect at the mean of select variables. The estimated marginal effects of the concentration variables shows that an increase in the dollars being serviced by Fannie servicers decreases the probabilities of modification and foreclosure and an increase in the dollars being serviced by all servicers in the MSA, increases the probability of modification and decreases the probability of foreclosure. The reduction in the probability of foreclosure is consistent with the idea that increased loan concentration leads servicers/lenders to consider the price consequence of foreclosure thus reducing their desire to foreclose, and instead to take no action.

As expected, some of the individual loan and property traits, as well as economic factors, impact the servicer's decision. As the current LTV ratio increases, the probability of foreclosure increases while the probability of modification and no action decreases. Loans with higher levels of mortgage insurance coverage have higher probabilities of foreclosure and modification. The modification marginal effect for mortgage insurance is over four times greater than for foreclosure marginal effect. Loans with longer period of time since their first severe delinquency (90 days+) have a higher probability of foreclosure and of no action being taken. This seems to reflect that once loans are delinquent beyond a certain length of time, they are left uncured. As interest rates increase, as measured by the change in the 10 year Treasury rate over the previous six months, the probabilities of modification and foreclosure decrease. However, an increase in interest rates since the time the loan was originated leads to a higher probability of modification and foreclosure. The marginal effect for the change in interest rates in the modification model is much larger than the marginal effect for the change in interest rates in the foreclosure model. This seems to imply servicers are attempting to negotiate new terms under the current market conditions.

Turning our attention back to the concentration of loans, a servicer's decision with regard to delinquent loans may not only be impacted by the local geographical concentration but also by the servicer's own national concentration. To account for the servicer's national portfolio concentration, we include the dollar amount of mortgages serviced by Fannie servicers nationally and all servicers across 20 MSAs in the model. The results are presented in Table 1.4. The coefficients of the variables in the base model, eq. (10), are relatively stable in sign and significance with the exception of the coefficient on our loan concentration variable – the dollar amount of loans serviced by all servicers in an MSA *m*. In the modification model, the sign of this coefficient switched from positive to negative, indicating that an increase in the dollar amount of all loans serviced in an MSA leads to a decrease in the probability of modification instead of the increase as predicted in the prior model. An increase in local loan concentration of Fannie loans leads to a lower probability of modification and foreclosure. The reduction in the probabilities of modification and foreclosure and corresponding increase in the probability of no action is consistent with the idea that an increase in the number of loans serviced hinders the servicer's ability to manage delinquent loans potentially due to lack of adequate staff and technology (Cordell et al (2008)). As the dollar amount of loans serviced across 20 MSAs by All servicers and nationally by Fannie servicers increases, the probabilities of modification and foreclosure increase. In the case of loans serviced a nationally by Fannie servicers the probability of a delinquent loan being modified is significantly greater than the probability of a delinquent loan being foreclosed, while for loans serviced across 20 MSAs by All servicers the probability of a delinquent loan being foreclosed is significantly greater than the probability of a delinquent loan being modified.

Thus far in the model, we have not considered the impact of foreclosure on the servicer's decision. To incorporate foreclosure into the model, we add the dollar amount of loans foreclosed by all servicers and Fannie servicers in the MSA. This model's estimates are presented in Table 1.5. Even with the addition of the foreclosure variable, the overall results stay the same, with the exception of the dollar amount of loans serviced by Fannie servicers nationally in the foreclosure part of the model. The coefficient on this variable changed from positive to negative, a result consistent with our expectations. An increase in the loans serviced by Fannie servicers nationally leads to a higher probability of modification and a lower probability of foreclosure. As previously stated, Fannie Mae holds the credit risk on all its loans and to minimize credit losses from defaulting loans, it maintains master servicing agreements to manage servicer behavior. When the number of number of loans serviced by Fannie servicer increases nationally, Fannie Mae is concerned about potential credit losses and directs servicers through their servicing agreements to be inclined to modify delinquent loans versus foreclosing on them. The coefficients on the new variable representing foreclosures of loans serviced by Fannie servicers is positive indicating that an increase in foreclosure activity increases the probability of foreclosure, as well as modification. This is consistent with the expected outcome. As servicers start foreclosing on loans, it will drive others to start foreclosure as not to be the only one left with a devalued portfolio.

When considering the impact of the servicers' decision on foreclosure, one has to consider the action of the other servicers in the MSA. In table 1.5 the variables representing foreclosure within the MSA are the sum of loans foreclosed by servicer *s,* in MSA *m;* loan foreclosed by *other Fannie* servicers and *non Fannie* loans foreclosed by *non-Fannie* servicers. To separate the effects of these variables we estimate the next model by substituting the loans foreclosed by all servicers in MSA *m* with *non-Fannie* loans foreclosed by servicers in MSA *m* (other foreclosures by all servicers) and loans foreclosed by Fannie servicers with *non-s* Fannie loans foreclosed by servicers in MSA *m* (Fannie foreclosures by other Fannie servicers). With this change, the overall results in Table 1.6 remain the same, with the exception of the dollar amount of loans serviced by Fannie servicers nationally in the foreclosure part of the model but it is statistically insignificant.

To summarize, as the concentration of Fannie loans in an MSA increases, the probability the Fannie servicer forecloses on a loan decreases, while the probability of modification decreases too, it is approximately 7 times smaller in magnitude than the decrease in the probability of foreclosure. This result is consistent with Fannie servicers internalizing the price impact of their own foreclosure actions. As foreclosures of *non-Fannie* servicers' loans increase in the MSA, servicer *s* is less likely to foreclose or modify a loan, but the likelihood of foreclosure is greater than modification. As the dollar magnitude of foreclosures of *non-s* Fannie servicers' loans increases, servicer *s* is more likely to modify or foreclose on a loan – a race to the bottom.

1.7 Issues of Endogeneity:

At the outset, it appears that the servicer controls the amount of Fannie loans they service in a geographical area and is the source of endogeneity. It should be noted that lenders/sellers themselves, not servicers, determine the geographic concentration of Fannie loans. Lenders/sellers originate and aggregate Fannie loans based upon the underwriting guideless of Fannie (uniform and consistent across lenders and markets). Upon the sale of the pool of loans by the lender/seller to Fannie, the lender retains most if not all loans with their mortgage servicing entity. In addition, the loans are originated over three years prior to the loan going severely delinquent (the average age at delinquency is 40 months). Given our focus on severely delinquent loans, we believe the workout area within the mortgage servicing area of an organization is functionally distant from the mortgage origination side (which creates market share) and even more distant from the overall banking area where market share is more likely being measured. That being said, servicing entities do have the ability to buy and sell mortgages from other servicing entities approved by Fannie in the secondary market (after securitization and servicer has on boarded the mortgage). These transactions (few and far between) typically occur to divest mortgages in markets wherein the servicer doesn't have the capacity or competence to service the mortgage or add mortgages to expand their portfolio.

The other potential source of endogeneity within this study, most directly, is Fannie Mae's choice to aggressively pursue certain markets, which directly impacts the dollar magnitude of Fannie loans outstanding in a particular geographic area. To understand how Fannie Mae's choice introduces endogeneity, let us review how the US government allocates discretionary spending and how Fannie Mae operates its business. The United States House Committee on Appropriations has jurisdiction over discretionary spending legislation for funding the federal government. The work of the Appropriations Committee is done in its subcommittees. The approved discretionary spending provides an economic boost to the local economy where it is being earmarked. There is a large body of empirical literature, which has investigated the impact of government spending on consumption, investment, and output variables. Studies have also found evidence linking congressional seniority to government spending; they also, find that senior members of the House were able to use their positions to improve their state's economic performance (Levitt and Poterba, 1999).

Let us now turn our attention to Fannie Mae and how it operates its business. As incorporated, the GSEs – Fannie and Freddie – have two conflicting missions, which work at cross purposes. As per the government directive they are required to keep their mortgage rates low and provide support for affordable housing, while the private shareholders seek to increase profitability by all means. Under this conflicting business interest, Fannie and Freddie provided political and financial support to Congress in return for freedom to preserve their lucrative lending business (Calomiris and Wallison, 2009). Between 2000 and 2008 the GSEs contributed $14.6 million to campaign of Senators/Representatives, most of which who were on various committees of Congress. Thus Fannie not only managed the credit and interest risk of their portfolio but also the political risk to ensure the profitability of their business. Hence, we believe that the tenure of a House representative on the House Appropriation Committee from the MSA in conjunction with political lobbying by Fannie has an impact on the amount of Fannie loans outstanding in an MSA. Hence, we believe that amount of loans serviced by Fannie servicers in an MSA might be the source of endogeneity in our estimating equation.

Given that lenders have the ability to choose programs they want to lend into and servicers have the ability to service loans in markets they choose, we posit that Fannie lenders will lend in to MSAs that exhibit robust economic growth and generate the highest profit. Hence, we use the tenure of House representatives on the House Appropriations Committee from the MSA as an instrument to predict the total number of Fannie loans serviced in an MSA. Intuitively, the House representatives with the longest tenures (senior most representatives) will be well positioned to garner funding for their constituency and coupled with Fannie Mae's political contributions to protect and grow their business, this instrument (tenure of the US Representative) predicts a larger number of Fannie mortgage loans originated in that MSA.

Within the MNL framework, we employ a control function estimation method to minimize the endogeneity bias, as suggested by Terza, Basu & Rathouz (2008) and was first proposed by Hausman (1978). The technique is a two-stage residual inclusion (2SRI) estimation framework. In the first-stage, a regression is estimated by regressing the endogenous variable on the exogenous variable. In the second-stage regression, the first-stage residuals are included as additional regressors in the second-stage estimation.

To operationalize the control function approach, we regress all Fannie mortgages managed by all servicers in a geographical area (all Fannie loans outstanding in an MSA) on all of the other explanatory variables contained in the MNL model in addition to a new variable representing the tenure of Congressional representative from the MSA on the House Appropriations Committee.

The control function approach has little overall impact on the estimates (in magnitude and significance) relative to the prior estimates. The estimates of the control function approach are presented in Table 1.7. In the first stage, the coefficient on the tenure of the House representative on the Appropriations Committee is positive and significant, as are a vast majority of the seller fixed effects. In the second stage, the coefficients and estimated marginal effects of the loan concentration variables remained consistent in sign, but increase substantially in magnitude.

Until this point we have evaluated the behavior of Fannie servicers as a group, in table 1.8, we specifically execute a 2SRI using the Fannie loans serviced by a servicer *s* in MSA *m*. In the first stage, we obtain residuals using the specification in table 1.7. In the second stage, we use the residual from first stage as a variable. We replace the variables representing the variable Fannie loans serviced in an MSA and loans serviced by All servicers in an MSA by Fannie loans serviced by a servicer in an MSA and Fannie loans serviced by a non-s servicer in an MSA (other servicers). The results, reported in Table 1.8, indicate that an increase in the dollar amount of mortgages serviced by a servicer in an MSA decreases the relative probability that a loan will be foreclosed vs. no action and decreases the relative probability that the delinquent loan will be modified vs. no action (the coefficient is statistically insignificant). The coefficients on the loans serviced nationally (portfolio concentration) are unchanged. Even with the substitution of variables, the overall results for the control variables stay the same.

In the two specifications using 2SRI, we find that servicers that service Fannie loans in an MSA – individually or collectively behave in the same manner with regard to managing delinquent loans. As the Fannie loans serviced by servicers in an MSA increase, the decrease in the probability of foreclosure is greater than the decrease in the probability of modification. Therefore as the concentration of Fannie loans in an MSA increase, the probability of foreclosure decreases.

1.8 Estimates of the Fixed Effect Variables:

The estimates of the fixed effect variables in the models estimated have been stable across the various models and contain some interesting insights outside of our primary hypotheses. The coefficients on the calendar time FE are one such example. From 2002 - 2007, the FE coefficients (all significant) were positive in both the modification and foreclosure models. However, the coefficients in the foreclosure model were a minimum of two times the magnitude of those in the modification model. In 2008, the coefficient in the modification model turned negative (figure 2). From 2010 onward, the calendar year FE coefficients in the modification model remained positive, but are now 1.3 to 5.4 times greater than the coefficient in the foreclosure models, where before they had been less. These results seem to show an impact of the GSE's 2008 decision to start paying servicers an incentive fee for each successful modification. Further, it indicates that servicers have some discretion in managing delinquent loans even when under a servicing contract with Fannie Mae.

The above results are in line with the timelines in other papers that have evaluated the numerous modification programs set in motion during and after the crisis. Scharlemann and Shore (2016) and Hembre (2014) have shown that the Home Affordable Modification Program (HAMP) reduced and prevented mortgage defaults. Maturana (2016) has studied the causal effect of loan modifications on loan losses and finds that losses on modified loans are one third less than the average loan loss (35.8% less). They have all provided evidence that the various modification programs enacted by the government have worked to a certain extent if not fully as intended.

Conklin et al (2018) find that seller-servicer affiliation increases the likelihood of modification. In the current study, loans originated through the seller's retail channel as opposed to the seller's correspondent channel are more (less) likely to be modified (foreclosed). Loans originated through the seller's retail channel are likely made to customers that have an existing relationship with the seller and, as a result, are more likely to be modified. Loans originated through the seller's broker channel, as opposed to the seller's correspondent channel, are more likely to be foreclosed. Interestingly, servicers are more likely to foreclose on loans in which they are identified as the seller than loans from other sellers.

Finally, given delinquency, loans underwritten by second homes and investment properties, as opposed to a primary residence, are less likely to undergo modification (more likely to be foreclosed upon) than a loan on a primary residence. This result is in line with S. Chan et al (2013). First time homebuyers are also more likely to be foreclosed on than non-first-time homebuyers. Loans on single family properties are more (less) likely to modified (foreclosed upon) than other property types in the data.

1.9 Conclusion:

The subprime crisis brought to the forefront the importance of mortgage servicing in managing delinquent loans. Previous studies concentrated on the difference in foreclosure and modification rates between securitized and non-securitized loans. The current study explores the decisions of agency (GSE) servicers to modify, foreclose, or take no action on delinquent loans. More importantly, it explores if agency mortgage servicers, as representative agents of investors/guarantor, internalize the potential negative price impact of foreclosure, through their servicing actions, to minimize the impact on the remaining pooled assets (under the direction of Fannie Mae).

Favara and Gianetti (2017) show that portfolio lenders, who have a large concentration of loans in a geographical area, will foreclose less often to minimize the decline in local house prices associated with foreclosure on their remaining portfolio. Our results, with respect to agency (GSE) servicers, are in line with the actions of portfolio servicers. We find that conditional on delinquency, an increase in the dollar amount of mortgages serviced by a Fannie servicer in an MSA decreases the probability that the delinquent loan will be foreclosed. The larger the concentration of loans, the lower the probability of foreclosing; this is an indication that Fannie servicers may be internalizing the potential for a negative price effect.

Foreclosures also play a role in the servicer's decision. As foreclosure activity of other servicers in the area increases, a servicer's probability of foreclosure increases. If other servicers' (*non-s*) foreclosure activity increases leading to a fall in the house prices, servicer (*s*) will increase foreclosure activity to reduce the value loss - foreclosures incentivize foreclosures.

A control function approach (a two-stage residual inclusion estimation framework) is used to reduce endogeneity bias that may exist in the model. The servicer's decision to foreclose a loan obviously has an impact, albeit small, on the total concentration of loans in a particular MSA. However, it is important to consider that the Fannie Mae actively manages its political risk (political lobbying and contributions) and its actions impact the cumulative amount of Fannie loans existing in a particular MSA. The control function estimation impacts the magnitude of the coefficient on the variable representing the dollar amount of Fannie loans in an MSA, but the coefficient remains consistent in sign.

Our results from the calendar time fixed effects shed some interesting insights, although not directly related to our main hypothesis. During the period 2002-2007, the coefficients from modification and foreclosure are both positive with the relative probability for foreclosure being twice as large as modification. In 2008, after federal and state government intervention (moratorium on foreclosures) and the GSE's decision to start paying servicers an incentive fee for each successful modification, there was a sharp reduction in the probability of foreclosure and a relative increase in the probability of modification.

The servicer decision with regard to delinquent loans is also influenced by loan and property characteristics as well as prevailing economic factors. The probability of foreclosure increases as the current LTV increases or the length of the delinquency period increases. For loans that have purchased mortgage insurance, the modification probability is three times the foreclosure probability. With regard to interest rate changes, the marginal effect for the modification outcome is much larger than the marginal effect for the foreclosure outcome, which implies servicers make an attempt to renegotiate loans under current market conditions.

The origination channel of the loan also plays a role in the servicer's decision. Loans originated through the seller's retail channel are more likely to be modified, as they are likely made to customers that have an existing relationship with the seller. Surprisingly, GSE servicers are more likely to foreclose on loans in which they are identified as the seller than loans from other sellers. Finally, given delinquency, loans to finance second homes and investment properties, as opposed to a primary residence, are more likely to be foreclosed upon. First time homebuyers are also more likely to be foreclosed on than non-first-time homebuyers.

Overall, the results indicate that agency (GSE) mortgage servicers behave in a manner consistent with prior research on portfolio servicers. The concentration of loans in a geographic area influences the decision to modify or foreclose on loans and the decisions are likely driven by the negative impact foreclosure has on the value of the remaining portfolio due to neighborhood price declines.

1.10 References:

Adelino, M., Gerardi, K., & Willen, P. S. (2013). Why don't Lenders renegotiate more home mortgages? Redefaults, self-cures and securitization. *Journal of Monetary Economics*, *60*, 835–853.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., Piskorski, T., & Seru, A. (2017). Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program. *Journal of Political Economy*, *125*(3), 654–712.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., & Evanoff, D. D. (2011). The role of securitization in mortgage renegotiation. *Journal of Financial Economics*, *102*, 559–578.

Anenberg, Elliot Kung, Edward. (2014). Estimates of the Size and Source of Price Declines Due to Nearby Foreclosures. *The American Economic Review, 104 (8),* 2527–51.

Been, V., Weselcouch, M., Voicu, I., & Murff, S. (2013). Determinants of the incidence of U.S. Mortgage Loan Modifications. *Journal of Banking and Finance*, *37*, 3951–3973.

Calomiris, C. W., & Wallison, P. J. (2009). The Last Trillion-Dollar Commitment: The Destruction of Fannie Mae and Freddie Mac. JOURNAL OF STRUCTURED FINANCE, (1), 71.

Campbell, J., Giglio, S., & Pathak, P. (2011). Forced Sales and House Prices. *The American Economic Review, 101*(5), 2108–2131.

Cordell, L., Dynan, K., Lehnert, A., Liang, N., & Mauskopf, E. (2008). Incentives of Mortgage Servicers: Myths and Realities. *Uniform Commercial Code Law Journal*, (4), 347–374.

Favara, G., & Giannetti, M. (2017). Forced Asset Sales and the Concentration of Outstanding Debt: Evidence from the Mortgage Market. *Journal of Finance, 72(3),* 1081–1118.

Foote, C., Gerardi, K., Goette, L., & Willen, P. (2010). Reducing Foreclosures: No Easy Answers. *NBER Macroeconomics Annual, 24*(1), 89–138.

Garriga, Carlos and Hedlund, Aaron, Mortgage Debt, Consumption, and Illiquid Housing Markets in the Great Recession (2017-10-01). FRB St. Louis Working Paper No. 2017–30.

Ghent, A. C. (2011). Securitization and Mortgage Renegotiation: Evidence from the Great Depression. *Review of Financial Studies*, *24*(6), 1814–1847.

Harding, J. P., Rosenblatt, E., & Yao, V. W. (2009). The contagion effect of foreclosed properties. *Journal of Urban Economics*, *66*, 164–178.

Hartley, D. (2014). The effect of foreclosures on nearby housing prices: Supply or dis-amenity? *Regional Science and Urban Economics*, *49,* 108–117.

Hembre, E. (2014, August). *Essays in Housing and Public Economics* (Ph.D.). University of Wisconsin.

Immergluck, D., & Smith, G. (2006). The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values. *Housing Policy Debate*, *17*(1), 57–79.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*, 305–360.

Samuel Kruger, The effect of mortgage securitization on foreclosure and modification, *Journal of Financial Economics*, 129, 586–607.

Levitin, A. J., & Twomey, T. (2011). Mortgage Servicing. *Yale Journal on Regulation*, *28*(1), 1–90.

Levitt, S., & Snyder, J. (1995). Political Parties and the Distribution of Federal Outlays. *American Journal of Political Science, 39*(4), 958-980.

Lin, Z., Rosenblatt, E., & Yao, V. (2009). Spillover Effects of Foreclosures on Neighborhood Property Values. *Journal of Real Estate Finance & Economics*, *38*(4), 387–407.

Maturana, G. (2017). When Are Modifications of Securitized Loans Beneficial to Investors? *Review of Financial Studies*, *30*(11), 3824–3857.

McQuade, Timothy and Guren, Adam, (2018), How Do Foreclosures Exacerbate Housing Downturns? *R&R, Review of Economic Studies*.

Mian, A., Sufi, A. and Trebbi, F. (2015), Foreclosures, House Prices, and the Real Economy. *The Journal of Finance, 70: 2587–2634*.

Piskorski, T., Seru, A., & Vig, V. (2010). Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis. *Journal of Financial Economics*, *97*(The 2007-8 financial crisis: Lessons from corporate finance), 369–397.

Posner, E. A. and L. Zingales. 2009. A loan modification approach to the housing crisis. *American Law and Economics Review* 11:575–607.

Ronel, E., Nicholas S., S., Souphala, C., Dennis, G., & Robert, H. (2010). What "Triggers" Mortgage Default? *The American Economic Review*, (2), 490–494.

Scharlemann, T. C., & Shore, S. H. (2016). The Effect of Negative Equity on Mortgage Default: Evidence From HAMP's Principal Reduction Alternative. *Review of Financial Studies*, *29*(10), 2850–2883.

Shleifer, A., & Vishny, R. (1992). Liquidation Values and Debt Capacity: A Market Equilibrium Approach. The Journal of Finance, 47(4), 1343–1366.

Zhu, J., Janowiak, J., Ji, L., Karamon, K., & McManus, D. (2015). The Effect of Mortgage Payment Reduction on Default: Evidence from the Home Affordable Refinance Program. *Real Estate Economics*, *43*(4), 1035–1054.

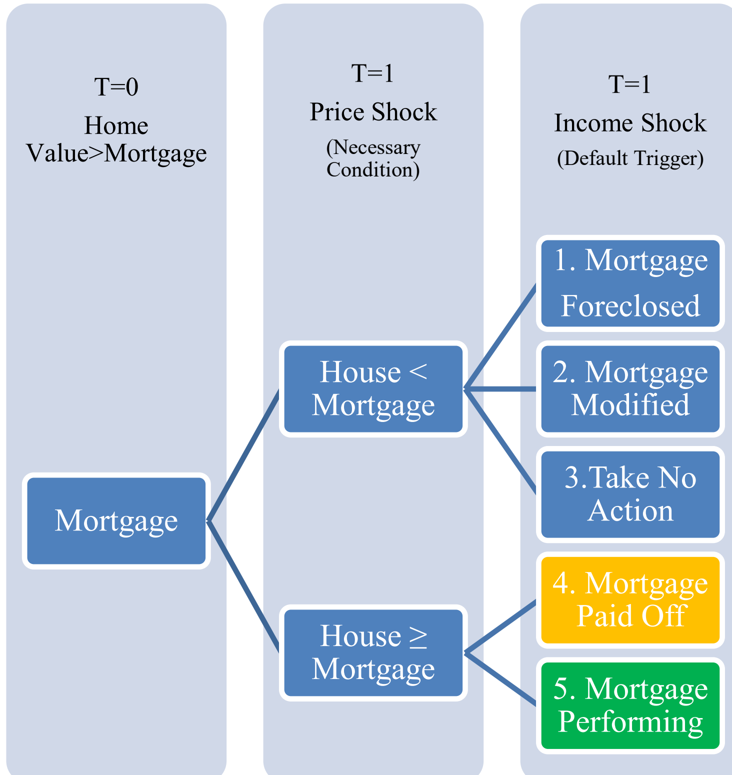

Figure 1.1: Dual shock model explaining mortgage delinquency.

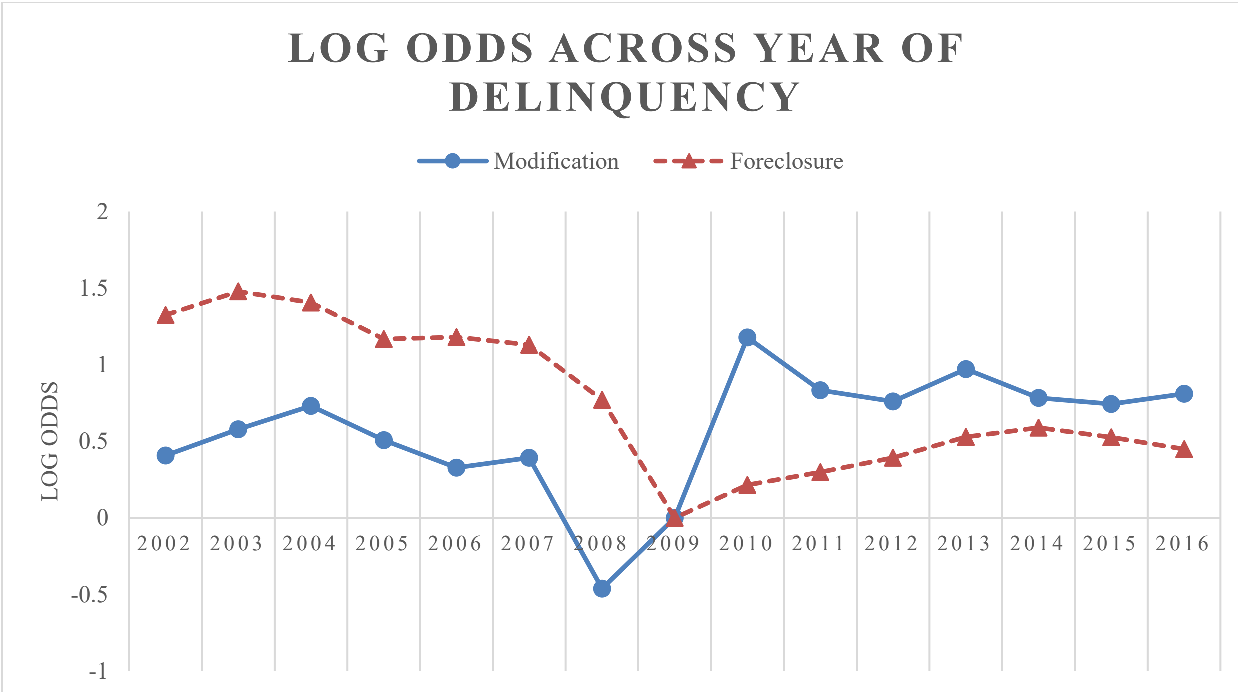

Figure 1.2: Log odds of foreclosure and modification across the sample period.

Table 1.1: Descriptive Statistics for the Pool of Delinquent Mortgage Loans

This table outlines the descriptive statistics of delinquent loans and the three categories based upon the loan outcomes – uncured, modified and foreclosed. The sample in this study consists of securitized first-lien agency mortgage loans that were at least 90-days delinquent. Mean values of the variables have been reported and the standard deviations are in parentheses below. (0,1) denotes dummy variable coded as 1=yes and 0=no.

Variable	Delinquent Loans	Uncured Loans	Modified Loans	Foreclosed Loans
At Origination				
Contract Rate	0.0615	0.0607	0.0611	0.0625
	(0.0061)	(0.0066)	(0.0057)	(0.0056)
Loan Amount	$211,225.78	$215,955.60	$235,908.52	$194,673.71
	($96,781.89)	($98,821.42)	($95,911.68)	($91,913.08)
Loan-to-value	0.7724	0.7464	0.7625	0.8042
	(0.1312)	(0.1412)	(0.1291)	(0.1136)
Debt to income ratio	0.4067	0.4036	0.4214	0.4029
	(0.1163)	(0.1177)	(0.1113)	(0.1166)
Mortgage Insurance (if required)	0.2414	0.2413	0.2384	0.2425
	(0.0644)	(0.0652)	(0.0661)	(0.0633)
At Delinquency				
Loan amount outstanding	$199,316.16	$202,990.75	$222,344.73	$184,495.68
	($93,860.33)	($95,808.69)	($93,544.75)	($89,268.84)
Current Loan-to-Value (CLTV)	0.8978	0.7917	0.9078	1.0030
	(0.2943)	(0.2664)	(0.2753)	(0.2917)
Loan age (in months)	49.69	51.39	51.55	47.06
	(30.21)	(33.43)	(28.41)	(27.19)
Occupancy *(0,1)*:				
Primary Home	0.9066	0.9238	0.9595	0.8639
Secondary Home	0.0214	0.0179	0.0133	0.0288
Investor	0.0720	0.0582	0.0273	0.1073
Property Type *(0,1)*:				
Condominium	0.1260	0.1081	0.0978	0.1579
Co-operative	0.0055	0.0081	0.0041	0.0033
Mfd. Housing	0.0038	0.0034	0.0025	0.0048
PUD	0.1577	0.1343	0.1525	0.1844
Single Family	0.7070	0.7460	0.7431	0.6496
Origination Channel *(0,1):*				
Correspondent	0.4526	0.4305	0.4587	0.4727
Retail	0.3176	0.3396	0.3142	0.2966
Broker	0.2297	0.2300	0.2271	0.2307
Purpose *(0,1):*				
Purchase	0.3189	0.3035	0.2803	0.3531
Refinance with Cash Out	0.4378	0.4600	0.4844	0.3928
Refinance with No Cash Out	0.2433	0.2365	0.2353	0.2540
Observations	359,388	148,438	67,651	143,129

Table 1.2: The average concentration of loans by servicer across time by MSA.
The averages in the table are found by calculating each servicer's concentration in an MSA at each point in time, averaging these concentrations across servicers at each point in time, and then finding the average over all time periods.

			2002-06		2007-11		2012-15	
MSA	**Avg**	**StDev**	**Avg**	**StDev**	**Avg**	**StDev**	**Avg**	**StDev**
Atlanta	0.0767	0.0227	0.0954	0.0147	0.0837	0.0098	0.0496	0.0110
Boston	0.0687	0.0219	0.0864	0.0133	0.0765	0.0097	0.0419	0.0094
Charlotte	0.0761	0.0214	0.0926	0.0152	0.0834	0.0094	0.0511	0.0111
Chicago	0.0726	0.0204	0.0880	0.0145	0.0799	0.0098	0.0486	0.0094
Cleveland	0.0471	0.0141	0.0576	0.0142	0.0503	0.0063	0.0328	0.0056
Dallas	0.0702	0.0203	0.0843	0.0137	0.0788	0.0093	0.0462	0.0116
Denver	0.0734	0.0242	0.0925	0.0130	0.0830	0.0106	0.0431	0.0110
Detroit	0.0700	0.0212	0.0873	0.0128	0.0770	0.0109	0.0443	0.0086
Las Vegas	0.0748	0.0245	0.0952	0.0171	0.0820	0.0110	0.0457	0.0092
Los Angeles	0.0774	0.0255	0.1006	0.0159	0.0829	0.0093	0.0471	0.0119
Miami	0.0741	0.0223	0.0933	0.0157	0.0794	0.0103	0.0483	0.0095
Minneapolis	0.0665	0.0168	0.0775	0.0103	0.0748	0.0081	0.0463	0.0094
New York	0.0732	0.0196	0.0894	0.0157	0.0776	0.0095	0.0515	0.0087
Phoenix	0.0721	0.0219	0.0902	0.0130	0.0796	0.0104	0.0453	0.0086
Portland	0.0740	0.0234	0.0937	0.0144	0.0812	0.0102	0.0457	0.0104
San Diego	0.0734	0.0227	0.0919	0.0120	0.0819	0.0104	0.0450	0.0097
San Francis.	0.0713	0.0217	0.0891	0.0135	0.0789	0.0089	0.0446	0.0084
Seattle	0.0659	0.0221	0.0854	0.0127	0.0721	0.0096	0.0388	0.0086
Tampa	0.0715	0.0209	0.0886	0.0148	0.0775	0.0097	0.0472	0.0097
Washington	0.0711	0.0167	0.0829	0.0102	0.0782	0.0087	0.0511	0.0087
Avg.	0.0710		0.0881		0.0780		0.0457	

Table 1.3 Estimate of the Fannie Servicer's Decision with Local Loan Concentrations.

*The estimates below are based on the estimation of a multinomial logit and the marginal effect (at means) using a sample of mortgages that are 90+ days delinquent (359,388 loans and 5,613,437 month observations).The dependent variable is defined by the Fannie servicer's decision to modify or foreclose or take no action (the base category). Estimates followed by * and + are statistically different from zero with 0.01 and 0.05 significance levels, respectively. The t-value for each coefficient is reported in parentheses.*

	Multinomial Logit		**Marginal Effect (at means)**		
Variable	**Modify**	**Foreclose**	**No Action**	**Modify**	**Foreclose**
Loans Serviced by *all* Servicers in MSA *m* ($, trillions)	0.291* (2.68)	-2.856* (20.09)	0.0453* (14.12)	0.00632* (3.18)	-0.0516* (20.19)
Loans Serviced by *Fannie* Servicers in MSA *m* ($, trillions)	-2.707* (12.20)	-16.34* (38.86)	0.338* (40.00)	-0.0440* (10.82)	-0.294* (39.12)
Loans Serviced by *all* Servicers across 20 MSAs ($, trillions)	X	X	X	X	X
Loans Serviced by *Fannie* Servicers Nationally ($, trillions)	X	X	X	X	X
Loans Foreclosed by *all* Servicers in MSA *m* (#, thousands)	X	X	X	X	X
Loans Foreclosed by *Fannie* Servicers in MSA *m* (#, thousands)	X	X	X	X	X
Loan Age at Delinquency *(in months)*	0.00156+ (2.44)	0.00569* (9.31)	-0.000129* (8.06)	0.0000267* (2.27)	0.000102* (9.27)
Current Loan-to-Value (CLTV)	-0.404* (22.78)	0.867* (52.97)	-0.00808* (18.54)	-0.00770* (23.72)	0.0158* (53.56)
Change in treasury rate over last 6 mos. ($y_t - y_{t-6}$)	-0.0934* (13.11)	-0.0104 (1.53)	0.00186* (10.51)	-0.00171* (13.09)	-0.000155 (1.27)
Change in treasury rate from Origination ($y_t - y_0$)	0.105* (12.57)	0.0392* (4.90)	-0.00258* (12.35)	0.00191* (12.49)	0.000672* (4.65)
Months Since Delinquency	-0.0250* (115.96)	0.00619* (40.45)	0.000340* (74.69)	-0.000460* (125.29)	0.000120* (43.61)
Mortgage Insurance Percentage	0.00249* (7.55)	0.000660* (2.27)	-0.0000564* (7.10)	0.0000454* (7.52)	0.0000111* (2.11)
Intercept	-4.028* (39.04)	-3.805* (1.17)			
MSA Fixed Effects	Yes	Yes			
Property Type Fixed Effects	Yes	Yes			
Loan Type Fixed Effects	Yes	Yes			
Origination Year Fixed Effects	Yes	Yes			
Delinquency Year Fixed Effects	Yes	Yes			
Servicer Fixed Effects	Yes	Yes			

Table 1.4 Estimate of the Fannie Servicer's Decision with Local Loan Concentrations.
*The estimates below are based on the estimation of a multinomial logit and the marginal effect (at means) using a sample of mortgages that are 90+ days delinquent (359,388 loans and 5,613,437 month observations).The dependent variable is defined by the Fannie servicer's decision to modify or foreclose or take no action (the base category). Estimates followed by * and + are statistically different from zero with 0.01 and 0.05 significance levels, respectively. The t-value for each coefficient is reported in parentheses.*

	Multinomial Logit		Marginal Effect (at means)		
Variable	**Modify**	**Foreclose**	**No Action**	**Modify**	**Foreclose**
Loans Serviced by *all* Servicers in MSA *m* ($, trillions)	-0.387* 3.30)	-4.501* (28.84)	0.0862* (24.81)	-0.00554* (2.59)	-0.0807* (28.98)
Loans Serviced by *Fannie* Servicers in MSA *m* ($, trillions)	-2.688* (12.03)	-15.75* (37.07)	0.326* (38.30)	-0.0438* (10.72)	-0.282* (37.26)
Loans Serviced by *all* Servicers across 20 MSAs ($, trillions)	0.469* (19.94)	0.697* (32.20)	-0.0207* (35.99)	0.00833* (19.40)	0.0124* (31.86)
Loans Serviced by *Fannie* Servicers Nationally ($, trillions)	3.108* (19.71)	0.537* (3.95)	-0.0652* (17.46)	0.0566* (19.68)	0.00858* (3.52)
Loans Foreclosed by *all* Servicers in MSA *m* (#, thousands)	X	X	X	X	X
Loans Foreclosed by *Fannie* Servicers in MSA *m* (#, thousands)	X	X	X	X	X
Loan Age at Delinquency *(in months)*	-0.00492* (7.03)	0.00293* (4.37)	0.0000367+ (2.11)	-0.0000909* (7.12)	0.0000542* (4.52)
Current Loan-to-Value (CLTV)	-0.400* (22.58)	0.868* (52.95)	-0.00810* (18.63)	-0.00762* (23.51)	0.0157* (53.52)
Change in treasury rate over last 6 mos. ($y_t - y_{t-6}$)	-0.139* (18.93)	-0.0518* (7.34)	0.00341* (18.59)	-0.00253* (18.82)	-0.000883* (6.97)
Change in treasury rate from Origination ($y_t - y_0$)	0.0858* (10.24)	0.0265* (3.27)	-0.00200* (9.57)	0.00156* (10.19)	0.000446* (3.07)
Months Since Delinquency	-0.0252* (116.55)	0.00610* (39.80)	0.000344* (75.72)	-0.000462* (125.99)	0.000118* (42.96)
Mortgage Insurance Percentage	0.00249* (7.56)	0.000690+ (2.38)	-0.0000569* (7.18)	0.0000453* (7.53)	0.0000115+ (2.22)
Intercept	-8.617* (39.13)	-5.793* (27.40)			
MSA Fixed Effects	Yes	Yes			
Property Type Fixed Effects	Yes	Yes			
Loan Type Fixed Effects	Yes	Yes			
Origination Year Fixed Effects	Yes	Yes			
Delinquency Year Fixed Effects	Yes	Yes			
Servicer Fixed Effects	Yes	Yes			

Table 1.5 Estimate of the Fannie Servicer's Decision with Local Loan Concentrations.
*The estimates below are based on the estimation of a multinomial logit and the marginal effect (at means) using a sample of mortgages that are 90+ days delinquent (359,388 loans and 5,613,437 month observations).The dependent variable is defined by the Fannie servicer's decision to modify or foreclose or take no action (the base category). Estimates followed by * and + are statistically different from zero with 0.01 and 0.05 significance levels, respectively. The t-value for each coefficient is reported in parentheses.*

	Multinomial Logit		Marginal Effect (at means)		
Variable	**Modify**	**Foreclose**	**No Action**	**Modify**	**Foreclose**
Loans Serviced by *all* Servicers in MSA *m* ($, trillions)	-0.583*	-2.667*	0.0563*	-0.00978*	-0.0465*
	(4.68)	(17.61)	(16.26)	(4.29)	(17.57)
Loans Serviced by *Fannie* Servicers in MSA *m* ($, trillions)	-1.809*	-10.99*	0.221*	-0.0294*	-0.192*
	(7.91)	(28.12)	(28.00)	(7.04)	(28.22)
Loans Serviced by *all* Servicers across 20 MSAs ($, trillions)	0.444*	0.441*	-0.0156*	0.00798*	0.00757*
	(18.63)	(20.30)	(27.06)	(18.31)	(19.93)
Loans Serviced by *Fannie* Servicers Nationally ($, trillions)	3.061*	-0.298+	-0.0499*	0.0561*	-0.00623*
	(19.29)	(2.19)	(13.44)	(19.36)	(2.62)
Loans Foreclosed by *all* Servicers in MSA *m* (#, thousands)	-0.00224*	-0.00833*	0.000183*	-0.0000383*	-0.000145*
	(11.38)	(40.39)	(36.31)	(10.61)	(40.29)
Loans Foreclosed by *Fannie* Servicers in MSA *m* (#, thousands)	0.406*	2.633*	-0.0525*	0.00655*	0.0459*
	(10.97)	(88.66)	(61.47)	(9.68)	(87.72)
Loan Age at Delinquency *(in months)*	0.000946	0.00457*	0.00000642	-0.0000880*	0.0000816*
	(1.52)	(6.80)	(0.37)	(6.88)	(6.94)
Current Loan-to-Value (CLTV)	-0.407***	0.693*	-0.00465*	-0.00760*	0.0123*
	(23.98)	(41.47)	(10.60)	(22.95)	(42.02)
Change in treasury rate over last 6 mos. ($y_t - y_{t-6}$)	-0.0928***	-0.0682*	0.00362*	-0.00247*	-0.00115*
	(13.37)	(9.59)	(19.88)	(18.36)	(9.24)
Change in treasury rate from Origination ($y_t - y_0$)	0.110***	0.0146	-0.00174*	0.00151*	0.000228
	(13.44)	(1.80)	(8.39)	(9.86)	(1.61)
Months Since Delinquency	-0.0239***	0.00614*	0.000347*	-0.000463*	0.000116*
	(116.35)	(39.73)	(76.90)	(126.21)	(42.86)
Mortgage Insurance Percentage	0.00259***	0.00224*	-0.0000844*	0.0000460*	0.0000384*
	(8.21)	(7.68)	(10.72)	(7.61)	(7.52)
Intercept	-3.668***	-4.569*			
	(37.39)	(21.71)			
MSA Fixed Effects	Yes	Yes			
Property Type Fixed Effects	Yes	Yes			
Loan Type Fixed Effects	Yes	Yes			
Origination Year Fixed Effects	Yes	Yes			
Delinquency Year Fixed Effects	Yes	Yes			
Servicer Fixed Effects	Yes	Yes			

Table 1.6 Estimate of the Fannie Servicer's Decision with Local Loan Concentration.
*The estimates below are based on the estimation of a multinomial logit and the marginal effect (at means) using a sample of mortgages that are 90+ days delinquent (359,388 loans and 5,613,437 month observations).The dependent variable is defined by the Fannie servicer's decision to modify or foreclose or take no action (the base category). Estimates followed by * and + are statistically different from zero with 0.01 and 0.05 significance levels, respectively. The t-value for each coefficient is reported in parentheses.*

	Multinomial Logit		**Marginal Effect (at means)**		
Variable	**Modify**	**Foreclose**	**No Action**	**Modify**	**Foreclose**
Loans Serviced by *all* Servicers in MSA *m* ($, trillions)	-0.513* (4.13)	-3.803* (24.29)	0.0753* (21.26)	-0.00810* (3.57)	-0.0672* (24.30)
Loans Serviced by *Fannie* Servicers in MSA *m* ($, trillions)	-1.670* (7.30)	-11.23* (28.54)	0.225* (28.11)	-0.0268* (6.40)	-0.198* (28.64)
Loans Serviced by *all* Servicers across 20 MSAs ($, trillions)	0.433* (18.18)	0.537* (24.57)	-0.0171* (29.57)	0.00774* (17.78)	0.00937* (24.23)
Loans Serviced by *Fannie* Servicers Nationally ($, trillions)	3.028* (19.10)	0.161 (1.19)	-0.0571* (15.34)	0.0553* (19.12)	0.00184 (0.77)
Non-s Fannie loans Foreclosed by Servicers in MSA *m* (#, thousands)	0.626* (14.39)	2.120* (59.79)	-0.0481* (47.38)	0.0107* (13.50)	0.0374* (59.18)
Non-Fannie loans foreclosed by Servicers in MSA *m* (#, thousands)	-0.00226* (11.46)	-0.00871* (42.35)	0.000192* (37.86)	-0.0000384* (10.65)	-0.000154* (42.26)
Loan Age at Delinquency *(in months)*	-0.00470* (6.71)	0.00350* (5.22)	0.0000235 (1.35)	-0.0000871* (6.81)	0.0000636* (5.35)
Current Loan-to-Value (CLTV)	-0.409* (22.59)	0.791* (47.51)	-0.00643* (14.60)	-0.00774* (23.41)	0.0142* (48.09)
Change in treasury rate over last 6 mos. ($y_t - y_{t-6}$)	-0.135* (18.37)	-0.0560* (7.90)	0.00340* (18.60)	-0.00245* (18.25)	-0.000947* (7.54)
Change in treasury rate from Origination ($y_t - y_0$)	0.0832* (9.93)	0.0147 (1.82)	-0.00175* (8.40)	0.00152* (9.90)	0.000233 (1.63)
Months Since Delinquency	-0.0252* (116.79)	0.00605* (39.19)	0.000347* (76.66)	-0.000463* (126.24)	0.000116* (42.32)
Mortgage Insurance Percentage	0.00260* (7.87)	0.00140* (4.79)	-0.0000709* (8.97)	0.0000471* (7.79)	0.0000239* (4.62)
Intercept	-8.292* (37.54)	-4.874* (23.11)			
MSA Fixed Effects	Yes	Yes			
Property Type Fixed Effects	Yes	Yes			
Loan Type Fixed Effects	Yes	Yes			
Origination Year Fixed Effects	Yes	Yes			
Delinquency Year Fixed Effects	Yes	Yes			
Servicer Fixed Effects	Yes	Yes			

Table 1.7 Estimate of the Fannie Servicer's Decision with Local Loan Concentrations utilizing a Control Function (2SRI).
*The estimates below are based on the estimation of a multinomial logit and the marginal effect (at means) using a sample of mortgages that are 90+ days delinquent (359,388 loans and 5,613,437 month observations).The dependent variable is defined by the servicer's decision to modify or foreclose or take no action (the base category). Estimates followed by * and + are statistically different from zero with 0.01 and 0.05 significance levels, respectively. The t-value for each coefficient is reported in parentheses. The standard errors for the multinomial logit in Stage II are bootstrapped.*

	Stage I	**Stage II**		**Marginal Effect (at means)**		
Variable	**Estimate**	**Modification Estimate**	**Foreclosure Estimate**	**No Action**	**Modify**	**Foreclose**
Loans Serviced by Fannie Servicers in MSA *m* ($, trillions)	X	-4.694* (2.64)	-7.354* (4.60)	0.212* (5.00)	-0.0833* (2.62)	-0.129* (4.51)
Term of House Representative on Appropriations Committee	0.000384* (392.48)	X	X	X	X	X
Loans Serviced by all Servicers in MSA *m* ($, trillions)	0.0743* (394.50)	-0.290 (1.93)	-4.089* (22.55)	0.0763* (16.26)	-0.00392 (1.21)	-0.0724* (21.02)
Loans Serviced by all Servicers across 20 MSAs ($, trillions)	-0.00329* (95.50)	0.423* (15.63)	0.550* (21.60)	-0.0171* (28.86)	0.00755* (16.89)	0.00960* (24.15)
Loans Serviced by Fannie Servicers Nationally ($, trillions)	0.0276* (139.88)	3.115* (19.35)	0.0481 (0.33)	-0.0567* (14.49)	0.0569* (18.78)	-0.000195 (0.08)
Non-s Fannie loans Foreclosed by Servicers in MSA *m* (#, thousands)	-0.0171* (247.40)	0.573* (9.62)	2.185* (55.63)	-0.0483* (38.95)	0.00975* (10.11)	0.0385* (49.19)
Non-Fannie loans Foreclosed by Servicers in MSA *m* (#, thousands)	0.000168* (610.32)	-0.00179* (5.88)	-0.00924* (29.83)	0.000193* (24.19)	-0.0000297* (4.90)	-0.000163* (30.91)
Residual from Stage I	X	3.075 (1.68)	-4.023+ (2.52)	X	X	X
Loan Age at Delinquency (in months)	-0.00000316* (3.03)	-0.00471* (6.59)	0.00353* (5.33)	0.0000231 (1.33)	-0.0000873* (6.83)	0.0000642* (5.40)
Current Loan-to-Value (CLTV)	-0.000959* (35.35)	-0.411* (21.61)	0.795* (46.79)	-0.00644* (14.57)	-0.00779* (23.48)	0.0142* (48.13)
Change in treasury rate over last 6 mos. ($y_t - y_{t-6}$)	-0.000287* (26.12)	-0.136* (17.26)	-0.0548* (7.37)	0.00340* (18.53)	-0.00247* (18.33)	-0.000925* (7.35)
Change in treasury rate from Origination ($y_t - y_0$)	0.000109* (8.75)	0.0835* (10.46)	0.0142 (1.77)	-0.00175* (8.38)	0.00152* (9.93)	0.000224 (1.57)
Months Since Delinquency	-0.0000171* (66.97)	-0.0253* (112.21)	0.00611* (35.27)	0.000347* (75.91)	-0.000464* (125.72)	0.000117* (42.16)
Mortgage Insurance Percentage	0.0000144* (29.58)	0.00265* (8.70)	0.00134* (4.77)	-0.0000708* (8.93)	0.0000479* (7.91)	0.0000229* (4.42)
MSA Fixed Effects	Yes	Yes	Yes			
Property Type Fixed Effects	Yes	Yes	Yes			
Loan Type Fixed Effects	Yes	Yes	Yes			
Origination Year Fixed Effects	Yes	Yes	Yes			
Delinquency Year Fixed Effects	Yes	Yes	Yes			
Seller Fixed Effects	Yes	No	No			
Servicer Fixed Effects	Yes	Yes	Yes			

Table 1.8 Estimate of the Servicer's Decision with Local Loan Concentrations utilizing a Control Function (2SRI).

*The estimates below are based on the estimation of a multinomial logit and the marginal effect (at means) using a sample of mortgages that are 90+ days delinquent (359,388 loans and 5,613,437 month observations).The dependent variable is defined by the servicer's decision to modify or foreclose or take no action (the base category). Estimates followed by * and + are statistically different from zero with 0.01 and 0.05 significance levels, respectively. The t-value for each coefficient is reported in parentheses. The standard errors for the multinomial logit in Stage II are bootstrapped.*

	Stage II		Marginal Effect (at means)		
Variable	**Modification Estimate**	**Foreclosure Estimate**	**No Action**	**Modify**	**Foreclose**
Fannie Loans Serviced by Servicer s in MSA m ($, trillions)	-0.609 (-0.28)	-5.754* (-2.91)	0.111+ (2.27)	-0.00918 (-0.26)	-0.102* (-2.91)
Fannie Loans Serviced by non-s Servicers in MSA m ($, trillions)	-4.905* (-3.03)	-12.02* (-7.96)	0.297* (7.51)	-0.0854* (-2.90)	-0.212* (-7.90)
Non Fannie Loans Serviced by all Servicers in MSA m ($, trillions)	-0.366+ (-2.06)	-4.130* (-21.20)	0.0784* (16.71)	-0.00528 (-1.63)	-0.0731* (-21.21)
Fannie Loans Serviced by Servicer s Nationally ($, trillions)	1.092* (5.79)	0.107 (0.64)	-0.0214* (-4.77)	0.0199* (5.79)	0.00152 (0.52)
Fannie Loans Serviced by non-s Servicers Nationally ($, trillions)	3.928* (23.10)	0.673* (4.53)	-0.0820* (-20.37)	0.0714* (23.08)	0.0106* (4.03)
Non Fannie Loans Serviced by all Servicers across 20 MSAs ($,trillions)	0.431* (17.58)	0.552* (24.58)	-0.0173* (-29.14)	0.00767* (17.18)	0.00964* (24.24)
Fannie Loans Foreclosed by non-s Servicers in MSA m (#, thousands)	0.552* (10.44)	2.186* (49.45)	-0.0479* (-38.58)	0.00934* (9.68)	0.0386* (49.06)
Non Fannie Loans Foreclosed by all Servicers in MSA m (#, thousands)	-0.00199* (-5.99)	-0.00923* (-30.91)	0.000196* (24.63)	-0.0000331* (-5.48)	-0.000163* (-30.84)
Residual from Stage I	2.333 (1.33)	-4.109+ (-2.53)	X	X	X
Loan Age at Delinquency (in months)	-0.00579* (-8.27)	0.00339* (5.04)	0.0297 (0.70)	0.0439 (1.37)	-0.0736+ (-2.56)
Current Loan-to-Value (CLTV)	-0.400* (-22.00)	0.799* (47.76)	0.0000447* (2.58)	-0.000107* (-8.36)	0.0000620* (5.21)
Change in treasury rate over last 6 mos. ($y_t - y_{t-6}$)	-0.135* (-18.27)	-0.0548* (-7.71)	-0.00675* (-15.29)	-0.00755* (-22.82)	0.0143* (48.32)
Change in treasury rate from Origination ($y_t - y_0$)	0.0865* (10.31)	0.0150 (1.86)	0.00337* (18.39)	-0.00244* (-18.15)	-0.000926* (-7.36)
Months Since Delinquency	-0.0254* (-116.98)	0.00609* (38.90)	-0.00181* (-8.69)	0.00157* (10.28)	0.000237 (1.65)
Mortgage Insurance Percentage	0.00258* (7.79)	0.00131* (4.48)	0.000349* (76.40)	-0.000466* (-126.40)	0.000117* (42.01)
MSA Fixed Effects	Yes	Yes			
Property Type Fixed Effects	Yes	Yes			
Loan Type Fixed Effects	Yes	Yes			
Origination Year Fixed Effects	Yes	Yes			
Delinquency Year Fixed Effects	Yes	Yes			
Seller Fixed Effects	No	No			
Servicer Fixed Effects	Yes	Yes			

CHAPTER 2

SPATIAL MISMATCH AND AFFORDABLE HOUSING

2.1 Introduction

Beginning in the 1950s and 1960s, U.S. metropolitan areas experienced a rapid suburbanization of jobs[7]. As a result, America's inner cities experienced sharp declines as people and businesses moved out to the suburban locations. In response to these economic shifts, Kain (1968) hypothesized that the combination of employment suburbanization and housing market discrimination reduced employment opportunities for metropolitan black workers and negatively affected their labor market outcomes.

The Spatial Mismatch Hypothesis (SMH), as this phenomenon became known, suggests that reducing the spatial barriers between jobs and housing can improve the economic prospects of disadvantaged groups. Over the last several decades, a number of policies have attempted to reduce these barriers, including housing dispersion programs that move inner-city residents to the suburbs, economic development programs designed to bring jobs to the inner cities where low-income workers live, and transportation and reverse-commuting programs to help inner-city residents get more easily to suburban jobs (Blumenberg & Pierce (2014) Chapple (2006), Ihlanfeldt & Sjoquist (1998)).

Spatial mismatch has proved to be remarkably persistent (Martin (2001), Martin (2004)). This paper studies one possible explanation for this persistence. Specifically, this paper uses tract-level employment and housing data for 103 U.S. metropolitan statistical areas between 2000 and 2010 to analyze whether jobs are moving to areas where affordable housing is less available. If this is so, then lower-income workers will find it more difficult to move to the areas that are gaining jobs. This, in turn, could reduce the labor market success of disadvantaged workers. This paper also tests whether there is a

[7] Edward L. Glaeser and Matthew E. Kahn, "Decentralized Employment and the Transformation of the American City." Brookings-Wharton Papers on Urban Affairs 2 (2001): 1–63

relationship between the extent to which jobs move to areas with fewer affordable housing units and the labor market outcomes of potentially disadvantaged workers.

We test this by first computing a Spatial Mismatch Index (SMI) and a General Spatial Mismatch Index (GSMI) for 103 U.S. metropolitan statistical areas between 2000 and 2010. We find that on average both the SMI and GSMI have decreased between 2000 and 2010 for both rental and ownership units, i.e., between 2000 and 2010 a smaller percentage of the population will have to move in order to make the distribution of employment and population across a metropolitan area identical.

When comparing the results of the SMI and GSMI across three categories of affordability (affordable to households at 80%, 100% and 120% of the area median income) for both rental and ownership units, we find that the results depart for certain categories, i.e., spatial mismatch has decreased under SMI and increased under GSMI. We find that the distribution of jobs and the distribution of rental units at 80% of the area median income has increased between 2000 and 2010 under GSMI. For ownership units, the distribution of jobs and the distribution of ownership units across all three categories of affordability has increased between 2000 and 2010 under GSMI.

When decomposing the SMI and GSMI shift into the employment shift and housing shift for rental units, the employment shift was positive in both cases. A positive employment shift implies that the changes in the distribution of metropolitan employment moved jobs farther away from the distribution of affordable housing in the year 2000. In the case of housing shift computed using GSMI reveals that it was positive for units that are affordable to households at 80 percent of the area median income, implying that the distribution of affordable rental units diverged from the 2010 distribution of jobs. In the case of ownership units, we find that employment shift was negative across all categories, implying that the distribution of jobs converged to the 2000 distribution of ownership units. The housing shift reveals a different story across the SMI and GSMI. The GSMI housing shift reveals that even though it was negative for all ownership units, it was positive across the three categories of affordable housing units. This implies that affordable ownership units diverged away from the 2010 distribution of jobs.

In this paper, we analyze the impact of spatial mismatch as measured by GSMI on the labor market outcomes of workers from the following four vulnerable groups – Black males, Black females, Hispanic males, and Hispanic females. The paper finds that, on average, using levels in 2000 and 2010, the labor market outcomes of Black males and females were impacted by an increase in the spatial mismatch between affordable rental housing and jobs, while Hispanic males and females were unaffected. A one percentage point increase in the spatial mismatch between rental units and jobs increased the unemployment rate of black males by 0.193 percentage points for rental units affordable at 100 percent of area median income, by 0.212 percentage points for rental units affordable at 120 percent of area median income, and by 0.234 percentage points for 'All' rental units. A one percentage point increase in the spatial mismatch between rental units and jobs increased the unemployment rate of black females by 0.084 percentage points for rental units affordable at 120 percent of area median income and by 0.102 percentage points for 'All' rental units.

A one percentage point increase in the spatial mismatch between rental units and jobs decreased the employment to population ratios of black males by 0.229 percentage points for rental units affordable at 80 percent of area median income, by 0.488 percentage points for rental units affordable at 100 percent of area median income, by 0.526 percentage points for rental units affordable at 120 percent of area median income, and by 0.558 percentage points for 'All' rental units. A one percentage point increase in the spatial mismatch between rental units and jobs decreased the employment to population ratios of black females by 0.125 percentage points for rental units affordable at 80 percent of area median income, by 0.309 percentage points for rental units affordable at 100 percent of area median income, by 0.339 percentage points for rental units affordable at 120 percent of area median income, and by 0.361 percentage points for 'All' rental units.

We then evaluated the impact of the change in GSMI between 2000 and 2010 on the change in the labor outcomes between 2000 and 2010. The labor market outcomes for the four groups were either not impacted or improved with an increase in the spatial mismatch between affordable housing and jobs for Black (males and females) and Hispanic females. Hispanic males were not affected by the change in the spatial mismatch between affordable housing and jobs. This result is inconsistent with previous results and provides no evidence that spatial mismatch between affordable housing and jobs contributes to the persistence of poor labor market outcomes of Blacks and Hispanics.

Additionally, the paper finds that Black and Hispanic workers between 2000 and 2010 experienced a negative effect on their unemployment rate (decreased) or a positive effect employment to population ratio (increased) in metropolitan areas where jobs shifted away from areas with greater numbers of affordable rental units. We find a similar result when the distribution of affordable housing shifted away from the location of jobs. It appears that despite an increase in the spatial mismatch, the poor labor market outcomes do not persist and the workers adapt to the change in the mismatch between jobs and affordable housing.

This paper contributes to the voluminous spatial mismatch literature in two ways. First, to the authors' knowledge, this is the first paper to analyze whether there is a relationship between spatial mismatch and the availability of affordable housing. Second, the paper updates previous work using the spatial mismatch index (Martin (2001) and (2004)) by using tract-level data and a distance-weighted index. Census tracts are amongst the most useful geographies from which we can develop neighborhood demographic-economic characteristics. They are designed to be fairly homogeneous with respect to population characteristics, economic status, and living conditions and are demarcated to represent neighborhoods/ localities. These features of known boundaries, covering the complete U.S., nesting geographic hierarchy and the availability of extensive demographic-economic data make census tracts an excellent choice to examine MSA characteristics and change. The benefit of using a distance-weighted – General Spatial Mismatch Index (GSMI) – is that it captures the spatial interaction between census tracts within an MSA and the spatial interaction between census tracts decreases as distance increases.

2.2 Literature Review

Since the seminal work carried out by Kain in the 1960s, a steady stream of literature has analyzed the spatial mismatch phenomenon. The spatial distribution of jobs and housing contributed to relatively poor labor market outcomes for black workers. Kain (1968) posits that residential segregation limits the housing mobility of black residents and consequently reduces their access to suburbanized jobs. If so, poor employment outcomes result from a “spatial mismatch” between residential locations and appropriate jobs. The spatial mismatch hypothesis (SMH) identifies three spatial factors:

1. Housing market segregation has limited the housing choices of black residents and constrained them in the inner cities;
2. Suburbanization of employment – especially low-skill and low-wage jobs – has reduced the number of suitable jobs for black workers living in the inner city; and
3. Black residents living in the inner cities do not have affordable or efficient transportation to travel to suburban jobs.

Ellwood (1986) offered evidence that geographical mismatch does not explain the unemployment rate amongst Black youth, thereby providing evidence against the spatial mismatch hypothesis. The paper found a significant but small effect of job accessibility on youth unemployment rates in the city of Chicago. Kasarda (1988) argues that there is a skills mismatch; the job opportunities within inner cities require high level skills that many inner-city job seekers do not possess, which results in inferior labor market outcomes for those residents.

Interest in this topic was revived in the 1990s and studies provided support for the spatial mismatch hypothesis (see Holzer (1991), Ihlanfeldt (1994), Kain (1992), Moss and Tilly (1991) and Wheeler (1990)). These papers measure spatial mismatch and then determine the impact of the spatial mismatch on poor labor market outcomes (high unemployment and low wages).

Holzer (1991) reviewed evidence on the spatial mismatch hypothesis between 1970 and 1990. He confirmed that spatial mismatch is germane to rationalize the difference between employment outcomes of Black and White workers. Despite extensive difference of opinion on the spatial mismatch hypothesis, he was able to conclude the following: Decentralization of population and employment continues across US metropolitan areas; Housing segregation has been declining with Blacks trailing Hispanics and

Asians; Blacks in central-city areas have lesser access to employment opportunities than Blacks or Whites in the suburbs; Employed Blacks have higher commute times than employed Whites.

Wheeler (1990) reviewed numerous spatial mismatch studies and found six studies that support the spatial mismatch hypothesis and three studies that found no evidence of residence/neighborhood effects on employment outcomes in the labor market. Wheeler suggested better and effective policy responses to the spatial mismatch hypothesis are warranted and future studies must control for gender and race.

Ihlanfeldt (1994) evaluates empirical evidence supporting the spatial mismatch hypothesis and concludes that job accessibility caused by spatial barriers results in the labor market struggles of inner city minority workers. This is further compounded by the decentralization of jobs and the less educated minority workers continuing to reside in the inner-city. This results in a spatial mismatch between jobs and the Black population and might explain the rapid rise of poverty in the inner-city.

The support for the spatial mismatch hypothesis is not unanimous. Studies such as Carlson and Theodore (1997), Cervero, Rood and Appleyard (1999), Ellen and Turner (1997), Ellwood (1998) and Immergluck (1998) question whether the policy prescriptions derived from the spatial mismatch hypothesis are the best approaches to improving the labor market outcomes of disadvantaged central-city workers.

Carlson and Theodore (1997) using individual-level data examine the relationship between community characteristics and job opportunities for neighborhoods in the city of Chicago. They find that residents of Black neighborhoods have poor access to employment, whereas those residing in poor neighborhoods do not encounter lesser job opportunities. Their results suggest that all low-wage populations might not be "spatially constrained".

Cervero, Rood and Appleyard (1999) determine that job accessibility is not a major factor in explaining black unemployment across the San Francisco Bay area. Their analysis also shows that residents of low-income, inner-city neighborhoods suffer from the highest employment mismatch, which support Carlson and Theodore (1997).

Ellwood (1998) explored the spatial dimensions of the labor market in Chicago. He found that job accessibility matters very slightly on the labor market outcomes of Blacks and the differential between Black and White employment rates is not explained by residential proximity to jobs.

Immergluck's primary conclusion in his book titled Neighborhood jobs, race, and skills: Urban unemployment and commuting (1998) is that race plays a very important role in the labor market outcomes in the US. He examines the relationship between employment and spatial mismatch including worker skills mismatch and poor job accessibility. He finds that race is a significant and persistent barrier to employment of Black workers resulting in their high unemployment rates.

While much of the early SMH research focused on the impact on black central-city workers, the impact of spatial mismatch on a variety of other vulnerable groups has been studied. These groups include immigrants, unskilled workers, low-income population and welfare recipients, and other racial and ethnic minority groups such as Asians and Latinos (Blumenberg (2004), Fan et al. (2014), Painter et al. (2007), Stoll (1999)).

Blumenberg (2004) evaluates the travel behavior of welfare participants and more specifically low-income women with regard to welfare-to-work transport programs policies designed to connect central city welfare participants to job opportunities in the suburbs and reduce their commute times. The paper finds that the travel patterns of low-income single mothers are not consistent with the programs resulting in a policy mismatch between welfare recipients and their transportation needs.

Painter et al. (2007) measure the impact of space and job accessibility on various race/ethnic populations among both youth immigrants and minority and the children of immigrants in the city of Los Angeles. The results suggest that both space and race play a significant part in the likelihood that a youth will be employed. The paper also highlighted the different impact on first or second generation immigrants.

Stoll (1999) studies intra-metropolitan (Los Angeles) labor market outcomes among Blacks/ Latinos and establishes the significant effect of spatial residential location and job accessibility on the poor labor market outcomes of Blacks/Latinos as compared to Whites. The results indicate that, ceteris paribus, the search for jobs by Blacks and Latinos is far more extensive than Whites. The extensive spatial job search is associated with greater employment for Blacks and higher wages for Latinos than that experienced by their White counterparts.

Further, a large number of studies have examined the issue of spatial mismatch in cities outside the U.S., including European cities (Fieldhouse (1999), Musterd and De Winter (1998)), Asian cities (Kawabata and Shen (2006), Liu and Wang (2011)), and Australian cities (Dodson (2005)).

Using 1991 Census to explore the layout of minority unemployment in London Fieldhouse finds that spatial mismatch does not explain the unemployment of Asians in London. It does though, on a limited basis explain Black unemployment. Musterd and De Winter (1998), on the other hand, have tested the hypothesis of spatial policy and housing policy on population segregation across 10 European cities. They conclude that globalization and role of the state on income distributions rather than spatial segregation are the causes of socio-economic inequalities.

Kawabata and Shen (2006) examine inter-metropolitan and intra-metropolitan differences in job accessibility by the mode of commuting in Boston, Los Angeles, and Tokyo. They find that job accessibility is lower for users of public transport across all three cities, but job accessibility for public transport users in the two US cities is significantly lower than in Tokyo.

Liu and Wang (2011) find that job-housing mismatch is starting to emerge in the city of Beijing but the patterns and mechanisms of spatial mismatch are distinct from the US. In Beijing, center-city neighborhoods have greater job accessibility as opposed to the suburbs. Dodson (2005) investigated spatial mismatch between location of housing affordability and employment in the city of Melbourne. The paper does not find strong support of the existence of spatial mismatch between affordable housing and employment opportunities 4in Melbourne between 1996 and 2001. The second part of the paper looks at the connection between labor market outcomes and access to transport. They conclude that poor public transport services impose a financial cost on low wage workers due to the need of owning an automobile.

The more recent studies have used innovative spatial mismatch measures such as dissimilarity indices and accessibility measures (Houston (2005a), Hu (2015), Martin (2004)).

Houston (2005a) discusses methods to measuring spatial mismatch. They include the analysis of labor market impact of residential segregation; comparison of commuting times; comparison of earnings; and measures of job proximity.

Hu (2015) adopted Shen's (1998) job accessibility measure, which accounts for the spatial distribution of competing workers, as well as the spatial distribution of jobs. In that context, Hu estimates the spatial barriers between poor job seekers and jobs appropriate to their skill and education levels in Los Angeles. Using average job accessibility for all poor job seekers as the yardstick helps to determine whether inner-city poor face spatial mismatch.

However, Martin (2001) uses the dissimilarity index as a measure of spatial mismatch to compute a spatial mismatch index. The spatial mismatch index captures MSA-level spatial disparity between jobs and population, providing a single value for each MSA. This index does not specify spatial imbalances within an MSA and does not incorporate any spatial interaction due to the distance between jobs and population within an MSA.

Finally, issues such as transportation and skills disadvantages have been studied in combination with the spatial mismatch hypothesis (Fan (2012), Fan et al. (2010), Houston (2005b), Taylor and Ong (1995)).

Taylor and Ong (1995) find that race and income do not explain commute times and automobile ownership is crucial to finding and maintaining employment, which they coined “automobile mismatch”. Using data from the American Housing Survey they find that while commute distances between Blacks, Hispanics and Whites are converging, the commute times were not converging with Blacks displaying the highest commute times versus Hispanics and Whites.

Fan et al (2010) examine the impact of the opening of a commuter rail line in Minneapolis on job accessibility and wages of low income workers. They find that proximity to stations and rail – bus connections are associated with gains in access to all workers, which suggests the importance of an integrated transport network. Fan’s (2012) review article contributes to the literature on Spatial Mismatch by evaluating spatial mismatch mitigation strategies. He suggests that promoting automobile ownership among low income workers is relatively the most effective way to address spatial mismatch. He views spatial mismatch as a barrier to commute and the automobile provides dependable transportation even if the residence and place of work are distant.

As highlighted, there is consensus amongst academics with regard to the existence of Spatial Mismatch in metropolitan areas, but the debate on its causes and the policies to counter it are ongoing. This paper first analyzes whether jobs are moving to areas where the availability of affordable housing is lacking or deficient and then to test if there is a relationship between spatial mismatch and the availability of affordable housing.

2.3 SMI/GSMI Analysis

2.3.1 Introduction:

The analytical tool that is used to compare the distributions of jobs and affordable housing units is the *spatial mismatch index (SMI)*, which was first used in Martin (2001, 2004) to study spatial mismatch. The SMI is based on the well-known dissimilarity index used to study residential segregation (Duncan and Duncan (1955), Taeuber and Taeuber (1965), and Massey and Denton (1968)). The SMI measures the degree to which jobs and people are located in different areas by measuring the percentage of the population that would have to move in order to make the distribution of employment and population across a metropolitan area identical. SMI measures the disparity across geographical divisions within an MSA like counties, zip codes or census tracts. In this paper, the SMI measures the extent to which an MSA's geographical distributions of employment and affordable housing differ across census tracts. For the most part census tracts are unchanging from census to census, except to subdivide due to the growth in population or to merge as a result of a reduction in population. Census tracts allow us to measure employment and housing shifts uniformly across the U.S. The most appealing feature of census tracts is geographic detail combined with availability of decennial census data and richer demographics from the American Community Survey (ACS). The other advantage of using census tract level data within an MSA to measure employment and affordable housing shifts as opposed to counties ensures that no MSAs are eliminated from the data set, as some MSAs are single-county. This analysis is performed for both rental and ownership units and for households at a variety of income levels. The analysis is performed using tract-level data for 2000 and 2010 across 103 U.S. metropolitan areas.

2.3.2 Description of the SMI and GSMI

The SMI for metropolitan area j is calculated as

$$SMI_j = \frac{1}{2H_j}\sum_{i=1}^{N_j}\left|\left(\frac{e_{ij}}{E_j}\right)H_j - h_{ij}\right| \tag{1}$$

where h_{ij} is the number of affordable units in census tract i in metropolitan area j, N_j is the number of census tracts belonging to metropolitan area j, e_{ij} is the total employment of census tract i in metropolitan area j, E_j is the total employment of metropolitan area j, and H_j is the total number of affordable units in metropolitan area j.

The SMI ranges between 0 (evenly distributed requiring no one to move) and 1 (unevenly distributed requiring everyone to move). Multiplying the index value by 100 makes the computed value intuitive to interpret i.e. percentage of the population that would have to move in order to make the distribution of employment and population across a metropolitan area identical.

One drawback of the SMI is that it does not take into account the physical distance between the location of jobs and affordable housing units. For example, suppose all affordable housing units are located in one census tract while all jobs are located in another census tract. Computing the SMI using equation *(1)* will yield a value of 100 irrespective of the distance between the two census tracts.

Therefore, we also use an alternative measurement index to capture the spatial interaction between census tracts by using distance – decayed composite counts and assume that the spatial interaction between two census tracts decreases as distance increases, which is a General Spatial Mismatch Index (GSMI) developed by Wong 2005.

$$GSMI_j = \frac{1}{2}\sum_{i=1}^{N_j}\left|\frac{De_{ij}}{DE_j} - \frac{Dh_{ij}}{DH_j}\right| \tag{2}$$

$$De_{ij} = \sum_{i=1}^{N_j}\frac{e_{ij}}{d_{ij}^2 + 1} \qquad DE_{ij} = \sum_{i=1}^{N_j} De_{ij}$$

$$Dh_{ij} = \sum_{i=1}^{N_j}\frac{he_{ij}}{d_{ij}^2 + 1} \qquad DH_{ij} = \sum_{i=1}^{N_j} Dh_{ij}$$

where Dh_{ij} is the composite weighted number of affordable units in census tract i and its interacting units in metropolitan area j, N_j is the number of census tracts belonging to metropolitan area j, De_{ij} is the composite weighted total employment of census tract i in metropolitan area j, DE_j is the composite weighted total employment of metropolitan area j, and DH_j is the composite weighted number of number of affordable units in metropolitan area j . All census tracts i in the metropolitan area j are considered as interacting with each other based upon a distance decay function $\frac{1}{d_{ij}^2+1}$. d_{ij} represents the aerial distance between the centroid of census tract i and the centroid of the other census tracts belonging to metropolitan area j.

The GSMI incorporates the spatial interaction information by utilizing the concept of composite counts (Wong 1998), which treat different population groups (in our case affordable housing units and jobs) in neighboring census tracts as if they are in the same census tract. The magnitude and effect of the neighboring census tracts is determined using the distance decay function $\frac{1}{d_{ij}^2+1}$. This distance decay function utilizes the square of the distance between census tracts within an MSA in the denominator, which causes the spatial interaction to decrease as the distance of the census tract increases. As a result, this measure accounts for the potential spatial interaction between jobs and housing units among neighboring census tracts. In this technique, the composite housing unit counts Dh_{ij} and the composite job counts De_{ij}, are computed first for each census tract in the MSA to derive the total composite housing unit counts DH_{ij} and the total composite job counts DE_{ij}, before executing equation 2. The computation procedure for Dh_{ij} and De_{ij} is similar to the process of deriving spatial moving averages for each census tract in the MSA, but the moving window is always changing, as the distance between the centroids of the census tracts differ and the distance decay function reduces the spatial interaction as the distance between census tracts increases. As a result the housing units and jobs that are at a great distance from each other will not be considered accessible despite being in the MSA.

Like the SMI, the GSMI is bound between 0 and 1, as it is computed in an identical fashion to the SMI, but instead uses composite population counts derived using a distance decay function, which incorporates spatial interaction into the index.

Both spatial mismatch indices are calculated for each metropolitan area using 2000 and 2010 employment data from the Census Transportation Planning Package and the 2000 and 2010 rental and owner-occupied housing data from the U.S. Census Bureau. Three values are calculated for each group: a 2000 value, a 2010 value, and a value using 2010 employment locations and 2000 housing locations (referred to as the "Mixed SMI"). This final measure is used to isolate the impact of 2000–2010 employment shifts on residential proximity to jobs. It also makes it possible to estimate whether 2000–2010 housing unit shifts offset or reinforced the impact of employment shifts.

The data used in this paper to calculate the SMI and GSMI are drawn from three primary data sources for 2000 and 2010: employment data from the Census Transportation Planning Package (CTPP); housing unit data, population and race data from the U.S. Census Bureau's American FactFinder database; and the Median Family Income (MFI) estimates from the department of Housing and Urban Development (HUD). First, the 2000 total employment data (jobs) at the census tract level is drawn from the Census Transportation Planning Package (CTPP) 2000, which is derived from the long form in Census 2000. The total employment data at the census tract level for 2010 is drawn from the Census

Transportation Planning Package (CTPP) based on the 2006 – 2010, 5-year American Community Survey (ACS) Data. The 2010 total employment data is constructed from five annual samples.

The data set for this study contains all census tracts belonging to metropolitan statistical areas (MSAs) with a population greater than 500,000 according to the Office of Management and Budget's MSA definitions for the year 2010. This restriction results in a sample of 103 MSAs containing 606 counties and over 47,000 census tracts. [8,9] Both employment as well as housing data is available at the census tract level for 2000 and 2010. Hence, we have been able to compute the spatial mismatch index for all the MSAs in our sample based upon the census tracts within an MSA. The resulting sample includes 47,232 census tracts that belong to the 103 largest MSAs in the US.

The share of total employment in a census tract within an MSA was gathered from the Census Transportation Planning Package (CTPP) 2000. It contains place of work data summarizing worker characteristics and we are able to compute the number of jobs across all categories of industries (total employment) within each census tract of an MSA. To obtain the share of total employment in a census tract within an MSA, we divided the total employment of each census tract by the total jobs/employment within that MSA.

To compute the share of affordable housing units in a census tract within an MSA we first need to establish a definition for housing affordability. Once the level of affordability is determined we apply it to the median family income estimates to obtain the share of income spent on housing in each MSA. Using these estimates we arrive at the distribution of the total number of affordable rental and owner occupied units in each census tract within the MSA. The details of the methodology applied are outlined below.

A common gauge of housing affordability is the share of income spent on housing. HUD currently uses a 30 percent of income standard for the Section 8 (rental assistance) voucher and certificate programs, we use the same ratio to compute housing affordability. Thus we consider a housing unit to be affordable if its monthly housing costs are less than or equal to 30 percent of gross monthly income. This

[8] Excludes Honolulu, HI MSA.
[9] Excludes Broomfield County, CO as it was incorporated November 15, 2001.

percentage of income standard has the advantage of being easy to calculate and easy to comprehend. In addition, data for computing this measure are readily available from a number of sources and standard Census tabulations show the share of households paying various shares of income for rent. Because the measure is a ratio, it can be easily compared across areas and over time.

The housing data for the distribution of rental and owner-occupied housing units in the 103 MSAs is gathered from the U.S. Census Bureau's American FactFinder database. We obtain the Median Family Income (MFI) estimates for each metropolitan area from the department of Housing and Urban Development (HUD). We compute the maximum monthly housing expense a household can apportion from the MFI estimates set by HUD for each MSA for 2000 and 2010 using the maximum monthly housing expense of 30 percent. We calculate the number of rental and owner-occupied housing units that are affordable to households at 80 percent, 100 percent, and 120 percent of MFI.

First, the rental housing units' data at the census tract level for 2000 and 2010 is drawn, which contains the distribution of the number of units at different rental ranges. To estimate the number of rental units available in a census tract in an MSA, we calculate the number of rental units that are available for rent below the maximum monthly housing expense for a household in each census tract of each MSA. Accordingly, we are able to calculate the number of rental units available at 80 percent MFI, at 100 percent MFI, and at 120 percent MFI. It is important to note here that the data is grouped at the census tract level and we are assuming a uniform distribution of units within each category.

The owner occupied housing units' data at the census tract level for 2000 and 2010 is drawn, which contains the distribution of the number of units at different ranges of house values. In order to estimate the number of owner-occupied housing units in each census tract, which are affordable we need to translate the maximum monthly housing expense into the maximum house the household can afford. For owner-occupied housing units the monthly housing expenses includes the mortgage payment as well as property taxes and insurance. We first calculate the maximum monthly mortgage payments assuming a down payment for the mortgage, the mortgage rate and its term. We have assumed a 20 percent down payment and utilized the average annual rate of 8.05 percent in 2000 and 4.69 percent in 2010, which are obtained from the Freddie Mac Primary Mortgage Market Survey of the 30yr Fixed Rate Mortgage.

Utilizing the American Community Survey data, we find that in 2000, the median property tax paid in the US was $1,334 and the median value of all housing units was $111,800, which informs us that the median property tax rate in 2000 was 1.1932 percent. In 2010, the median property tax paid was $2,245 and the median value of all housing units was $188,400, for a median property tax of 1.1916 percent. Using these ratios, we have estimated the median property tax rate for 2000 and 2010 at 1.20 percent of house value. Zillow has estimated the average annual property insurance in the US at 0.42 percent of house value ($35 per month for every $100,000 of house value). People in high risk zones (areas prone to natural disasters, crime and other perils) can expect to pay a higher premium, as can people who add additional coverage to their policies (for things such as floods or personal property). Using an average or median for property taxes and property insurance premiums for the US as a whole are not ideal due to the wide variations in local economic conditions, real estate values and hazards from state to state. Therefore, to ensure that we err on the side of caution, we assume property taxes and insurance as a percentage of house value at 2.00 percent in 2000 and 2010. After applying these assumptions, we estimate the threshold value of a house in each MSA that can be purchased such that the maximum monthly housing expense is below 30 percent of MFI.

To estimate the number of owned units available in a census tract in an MSA, we calculate the number of owner-occupied units that are below the threshold value of a house in each census tract of each MSA (which was estimated such that maximum monthly housing expense for a household does not exceed 30 percent of MFI). Accordingly, we are able to calculate the number of owned units available at 80 percent MFI, at 100 percent MFI, and at 120 percent MFI. Here again, the data is grouped at the census tract level and we are assuming a uniform distribution of units within each category.

2.3.3 Results from the SMI and GSMI

Two sets of spatial mismatch indices are calculated for each of the 103 metropolitan areas using 2000 and 2010 population data from the U.S. Census Bureau's American FactFinder database and 2000 and 2010 employment data from Census Transportation Planning Package (CTPP) 2000 for each of the housing unit categories. The first set calculates the spatial mismatch index and the general spatial mismatch index for rental units within each MSA and the second set calculates the same indices for owned units within each MSA. Within each set, four separate indices are calculated for each metropolitan area: all units within the category, units at 80 percent of MFI within the category, units at 100 percent of MFI within the category, units at 120 percent of MFI within the category.

Three values are calculated for each index: a 2000 value, a 2010 value, and a value using 2010 employment locations and 2000 housing locations (referred to as the "mixed" SMI/GSMI). This final measure is used to isolate the impact of 2000–2010 employment shifts on residential unit proximity to jobs (the employment shift) and to predict the effect of changes in employment locations on the unemployment rates of the race–gender groups of workers in the study. The "mixed" SMI/GSMI also makes it possible to estimate the impact of 2000–2010 change in the distribution of affordable housing units on the relative locations of jobs by comparing the mixed value for a particular decade with the next ''non-mixed'' SMI (the housing shift). It is important to note that the housing shift is not changes or shifts in the physical units. In our context, it measures the impact of new affordable units or the changes in rents/values of existing units. In this paper, we cannot distinguish between these two effects. Tables 2.1 and 2.2 provide a summary of the weighted average SMIs and GSMIs across the 103 MSAs (the 103 MSAs in the sample have been weighted by their respective 2010 population).

Table 2.1 provides a summary of the weighted average SMIs and GSMIs for rental units across the 103 MSAs. We find that the spatial mismatch between rental housing and jobs decreased when using the SMI across all four categories ('All' units, 80 percent of MFI, 100 percent of MFI, 120 percent of MFI). The values presented are weighted averages according to the total MSA population in the ending year of the decade. For rental units the reduction in the SMI was greatest for units at 80 percent of MFI (change is

-4.80 percent) and the reduction decreased as the percentage of MFI increased. The reduction in the SMI for units at 100 percent of MFI is -4.27 percent and at 120 percent of MFI is -3.82 percent. The jobs – rental units SMI decreased in 73 MSAs and increased in 30 MSAs. Interestingly, when using the GSMI to measure the spatial mismatch between rental housing and jobs, we find that the GSMI decreased for all rental units (change is -3.48 percent) but the GSMI increased for the units at 80 percent of MFI (change is +6.74 percent), decreased for the units at 100 percent of MFI (change is -0.76 percent) and 120 percent of MFI (change is -2.65 percent). The jobs – rental units GSMI decreased in 58 MSAs and increased in 45 MSAs, which is 15 MSAs fewer (20.5 percent lesser) than the jobs – rental units SMI. For rental units, both the SMI and GSMI confirm that the spatial mismatch between rental units and jobs has decreased between 2000 and 2010 but the GSMI reveals that the distribution of rental units affordable to households with incomes equal to 80 percent of MFI has diverged from the distribution of jobs and converged at 100 percent and 120 percent of MFI.

Looking at the rows labeled "mixed SMI" and comparing them to the 2000 SMI values reveals the contribution of 2000 to 2010 employment shifts on the total shift, holding the housing distribution constant at its 2000 levels. In Table 2.1 (rental units), the employment shift SMI for residents within the MSA increased from 45.27 to 45.43 (0.35 percent) when the 2010 job distribution is used instead of the 2000 job distribution. This means that employment growth between 2000 and 2010 was higher in tracts with relatively fewer rental units and the distribution of jobs diverged from the distribution of rental housing. The employment shift is the greatest for rental units at 80 percent of MFI increasing from 44.43 to 44.70 (0.6 percent) and decreases as the MFI increases. The employment shift for rental units at 100 percent of MFI increases from 44.98 to 45.20 (0.47 percent) and at 120 percent of MFI increases from 45.17 to 45.35 (0.41 percent). This means that jobs have moved away from tracts with the most affordable units. We observe a similar pattern when comparing the "mixed GSMI" to the 2000 GSMI values. The employment shift is the greatest for rental units at 80 percent of MFI increasing from 13.12 to 13.25 (0.95 percent) and decreases as the MFI increases. The employment shift for rental units at 100 percent of MFI increased from 12.82 to 12.91 (0.73 percent) and at 120 percent of MFI increased from 12.77 to 12.85 (0.6 percent).

Comparing the "mixed SMI" values with the 2010 values reveals the impact of 2000 to 2010 in the distribution of affordable housing units on the total shifts, holding the employment distribution constant at its 2010 levels. In table 1 (rental units), the housing shift SMI for residents within the MSA decreased from 45.43 to 43.75 (-3.71 percent) when the 2010 housing distribution is used instead of the 2000 housing distribution. This means that distribution of rental housing converged to the distribution of jobs. The reduction in housing shift SMI is the greatest for rental units at 80 percent of MFI decreased from 44.70 to 42.30 (-5.37 percent) and decreases as the MFI increases. The housing shift for rental units at 100 percent of MFI decreased from 45.20 to 43.07 (-4.72 percent) and at 120 percent of MFI decreased from 45.35 to 43.44 (-4.22 percent).This may be interpreted that the distribution of rental units at 80 percent of MFI have converged more than rental units at or above the MFI. Just like the housing shift SMI, the housing shift GSMI for residents within the MSA decreased from 12.79 to 12.29 (-3.95 percent). Interestingly, when using the GSMI, we find that the housing shift GSMI decreased for all rental units but the housing shift GSMI for the units at 80 percent of MFI increased from 13.25 to 14.01 (5.74 percent) and decreases as the MFI increases. The housing shift for rental units 100 percent of MFI decreased from 12.91 to 12.72 (-1.48 percent) and at 120 percent of MFI decreased from 12.85 to 12.43 (-3.23 percent). Using a distance weighted measure, reveals that even though the change in distribution between rental units and jobs has decreased between 2000 and 2010, the distribution of units at the most affordable level (80 percent of MFI) have diverged from the distribution of jobs i.e. distance between units at the 80 percent of MFI and jobs has increased and the distance between units at the 100 percent and 120 percent of MFI has decreased less than the SMI.

Table 2.2 provides a summary of the weighted average SMIs and GSMIs for ownership units across the 103 MSAs. We find that the spatial mismatch between ownership housing and jobs decreased when using the SMI (change is -5.55 percent) across all four categories (all units, 80 percent of MFI, 100 percent of MFI, and 120 percent of MFI). The values presented are weighted averages according to the total MSA population in the ending year of the decade. For ownership units the reduction in the SMI was lowest for units at 80 percent of MFI (change is -3.14 percent) and the reduction increased as the percentage of MFI increased. The reduction in the SMI for units at 100 percent of MFI is -4.61 percent and at 120 percent of MFI is -5.76 percent. The jobs – owned units SMI decreased in 78 MSAs and increased in 25 MSAs. Interestingly, when using the GSMI to measure the spatial mismatch between ownership housing and jobs, we find that the GSMI decreased for all ownership units (change is -12.08 percent) but the GSMI increased for the units at 80 percent of MFI (change is 5.92 percent), increased for the units at 100 percent of MFI (change is 7.57 percent), increased for the units at 120 percent of MFI (change is 5.63 percent). Using a distance weighted measure, reveals that even though the change in the overall distribution between ownership units and jobs has decreased between 2000 and 2010, the distribution of units affordable to households at each of the three income thresholds has diverged from the distribution of jobs i.e. distance between units at the three affordable levels and jobs has increased. For ownership units, both the SMI and GSMI confirm that the spatial mismatch between ownership units and jobs has decreased between 2000 and 2010 but the GSMI reveals that the distribution of ownership units at 80 percent, 100 percent, and 120 percent of MFI has diverged from the distribution of jobs.

In Table 2.2, the employment shift SMI for residents within the MSA decreased from 49.94 to 48.53 (-2.82 percent) when the 2010 job distribution is used instead of the 2000 job distribution. This means that employment growth between 2000 and 2010 was higher in tracts with relatively greater ownership units and the distribution of jobs converged to the distribution of ownership housing. The employment shift is the smallest for ownership units at 80 percent of MFI decreasing from 43.64 to 43.20 (-1.01 percent) and increases as the MFI increases. The employment shift for ownership units at 100 percent of MFI decreased from 44.71 to 43.93 (-1.74 percent) and at 120 percent of MFI decreased from 45.87 to 44.88 (-2.15 percent). This means that jobs have moved closer to tracts with affordable units. We observe a similar pattern when comparing the “mixed GSMI” to the 2000 GSMI values. The employment shift is the smallest for ownership units at 80 percent of MFI decreasing from 26.52 to 26.37 (-0.59 percent) and the magnitude converges more as the MFI increases. The employment shift for ownership units at 100 percent of MFI decreased from 23.31 to 22.96 (-1.52 percent) and at 120 percent of MFI decreased from 21.97 to 21.49 (-2.19 percent). The jobs – owned units GSMI decreased in 50 MSAs and increased in 53 MSAs.

In Table 2.2, the housing shift SMI for residents within the MSA decreased from 48.53 to 47.17 (-2.81 percent) when the 2010 housing distribution is used instead of the 2000 housing distribution. This means that distribution of ownership housing converged to the distribution of jobs. The reduction in housing shift SMI is the smallest for ownership units at 80 percent of MFI decreased from 43.20 to 42.27 (-2.15) and the magnitude converges more as the MFI increases. The housing shift for ownership units at 100 percent of MFI decreased from 43.93 to 42.65 (-2.92 percent) and at 120 percent of MFI decreased from 44.88 to 43.22 (-3.69 percent). The jobs – owned units SMI decreased in 67 MSAs and increased in 36 MSAs. This may be interpreted that the distribution of owned units affordable to households with incomes equal to 80 percent of MFI have converged less than ownership units at or above the MFI. Just like the housing shift SMI, the housing shift GSMI for residents within the MSA decreased from 21.02 to 19.17 (-8.82 percent). Interestingly, when using the GSMI, we find that the housing shift GSMI for ownership units increased across all three categories of affordability (80 percent of MFI, 100 percent of MFI, and 120 percent of MFI). The employment shift for ownership units at 80 percent of MFI increased from 26.37 to 28.09 (6.54 percent) at 100 percent of MFI increased from 22.96 to 25.08 (9.22 percent) and at 120 percent of MFI increased from 21.49 to 23.21 (7.99 percent). The jobs – owned units SMI decreased in 34 MSAs and increased in 69 MSAs. Using a distance weighted measure, reveals that even though the change in the overall distribution between ownership units and jobs has decreased between 2000 and 2010, the distribution of units at the three levels of affordability diverged from the distribution of jobs i.e. distance between units at the three levels of affordability and jobs has increased.

2.3.4 Summary

Across the 103 MSAs, we find that the SMI and GSMI have decreased between 2000 and 2010 for both rental units and ownership units. The spatial mismatch for ownership units has decreased greater than rental units. Key differences begin to appear when the housing market at 80 percent of MFI (affordable to low income households), 100 percent of MFI and 120 percent of MFI are analyzed (affordable to moderate income households). When computing the spatial mismatch index using GSMI, we find that the distribution of rental units at 80 percent of MFI has diverged from the distribution of jobs i.e. spatial mismatch has increased between 2000 and 2010. This is even more evident when we evaluate ownership units. The distribution of ownership units at all three levels of affordability have diverged from the distribution of jobs i.e. spatial mismatch has increased between 2000 and 2010.

When decomposing the SMI/GSMI shift into the employment shift and housing shift for rental units, we find that employment shift was positive across all categories, implying that the growth in jobs occurred away from affordable rental units. Housing shift across all categories were negative (except for rental units at 80 percent of MFI), implying that the distribution of affordable rental units converged to the 2010 distribution of jobs. The GSMI housing shift reveals that the housing shift for rental units at 80 percent of MFI was positive, implying that the distribution of affordable rental units at 80 percent of MFI diverged from the 2010 distribution of jobs.

When decomposing the SMI/GSMI shift into the employment shift and housing shift for ownership units, we find that employment shift was negative across all categories, implying that the distribution of jobs converged to the distribution of ownership units. The housing shift reveals a different story across the SMI and GSMI. We find that the SMI housing shift across all categories of ownership units was negative, implying that the distribution of affordable ownership units converged closer to the 2010 distribution of jobs. The GSMI housing shift reveals that even though it was negative for all ownership units, it was positive across the three categories of affordable housing units. This implies that ownership units at 80 percent, 100 percent, and 120 percent of MFI diverged away from the 2010 distribution of jobs.

2.4 Regression Analysis

2.4.1 Introduction:

In the previous section, we computed the SMI and GSMI across 103 MSAs between 2000 and 2010 and analyzed how the distribution between jobs and affordable housing units changed in that period. We further decomposed the total shift in SMI/GSMI into the employment shift SMI/GSMI and the housing shift SMI/GSMI. The second part of this paper focuses on whether the disparity between an MSA's geographical distributions of employment and affordable housing units affects the labor market performance of the four vulnerable groups (Black males, Black females, Hispanic males, and Hispanic females). In this stage of the analysis, two measures of labor market outcomes – unemployment rate and employment to population ratio – are regressed on variables representing key determinants of labor market outcomes and, most importantly, the spatial mismatch index values. Separate regressions are performed for each of the four vulnerable groups using the GSMI. The 2000 – 2010 change in the MSA – level unemployment rates and MSA – level employment to population ratios are computed to enable us to determine if the change in the distribution of jobs and housing units are related to the changes in the labor market outcomes of Black and Hispanic workers. If higher GSMI values are found to be related to higher unemployment rates (lower employment to population ratios) then it suggests that GSMI negatively affects the labor market outcomes of the vulnerable groups.

2.4.2 Presentation of regression models

To test whether GSMI or change in GSMI affects the labor market outcomes of the vulnerable groups, the following regression equations are estimated using ordinary least squares, with the observations weighted by their shares of total metropolitan population in 2010. The first model estimates the relationship using level values of the variables for 2000 and 2010 in the regression equation. The second model estimates the relationship using the change in the values of the variables between 2000 and 2010. The third and final regression, splits the change in GSMI into change in employment shift and a change in housing shift. This final specification enables us to determine if the impact of change in GSMI is a function of the jobs moving away from affordable housing or if the distribution of affordable housing is moving away from jobs.

i. Model 1: unemployment rate = f(GSMI levels along with the 2000 and 2010 values of control variables)

$$UNEMP\_RATE_{i,g,t} = \alpha_{i,t} + \beta_1 * TOT\_GSMI_{i,h,t} + \gamma * \boldsymbol{X} + \delta * CEN\_YEAR + \varepsilon_{i,t} \qquad (3)$$

$$EMP\_POP\_RATIO_{i,g,t} = \alpha_{i,t} + \beta_1 * TOT\_GSMI_{i,h,t} + \gamma * \boldsymbol{X} + \delta * CEN\_YEAR + \varepsilon_{i,t} \qquad (4)$$

	Dependent variable
$UNEMP\_RATE_{i,g,t}$	MSA level unemployment rate for Black/ Hispanic workers (males and females), where i represents the MSA, g the vulnerable group and t the census year.
$EMP\_POP\_RATIO_{i,g,t}$	MSA level employment to population ratio for Black/ Hispanic workers (males and females), where i represents the MSA, g the vulnerable group and t the census year.
	Primary variable of interest
$TOT\_GSMI_{i,h,t}$	MSA Total GSMI, where i represents the MSA, h the housing unit type and t the census year.
	Dummy variable
CEN_YEAR	A dummy variable, which takes the value 0 for the year 2000 census observations and the value 1 for the year 2010 census observations.

The labor market outcomes of the four vulnerable groups – Black males, Black females, Hispanic males, and Hispanic females is represented by the following 'level' variables in Model 1: $UNEMP\_RATE_{i,g,t}$ and $EMP\_POP\_RATIO_{i,g,t}$. The $UNEMP\_RATE_{i,g,t}$ represents the MSA level unemployment rate of the vulnerable group and $EMP\_POP\_RATIO_{i,g,t}$ represents the MSA level employment to population ratio of the vulnerable group.

$TOT\_GSMI_{i,h,t}$ is the primary variable of interest in this equation, which will enable us to evaluate the impact of the spatial mismatch level (GSMI) between jobs and housing units within an MSA on the labor market outcomes of vulnerable groups. $TOT\_GSMI_{i,h,t}$ represents the level of GSMI in an MSA, for a housing type, in a census year. In equation 1a, a positive coefficient on the spatial mismatch level (GSMI) can be viewed as evidence that the shift of job locations away from affordable housing locations negatively impacted labor market outcomes i.e. increases the unemployment rate of the vulnerable group. In equation 1b, a negative coefficient on the spatial mismatch level (GSMI) can be viewed as evidence that the shift of job locations away from affordable housing locations negatively impacted labor market outcomes i.e. decreases the employment to population ratio of the vulnerable group.

CEN_YEAR is a dummy variable and takes the value 0 for the year 2000 census observations and the value 1 for the year 2010 census observations. This allows us to control for the differences between the census year 2000 and 2010.

ii. Model 2: Change in unemployment rate = f(Change in GMI and the changes in control variables)

$$\Delta UNEMP\_RATE_{i,g} = \alpha_i + \beta_1 * GSMI\_EFF_{i,h} + \gamma * \boldsymbol{X} + \varepsilon_1 \tag{5}$$

$$\Delta EMP\_POP\_RATIO_{i,g} = \alpha_i + \beta_1 * GSMI\_EFF_{i,h} + \gamma * \boldsymbol{X} + \varepsilon_1 \tag{6}$$

	Dependent variable
$\Delta UNEMP\_RATE_{i,g}$	Change in MSA level unemployment rate between 2000 and 2010 for Black/ Hispanic workers (males and females), where i represents the MSA and g the vulnerable group.
$\Delta EMP\_POP\_RATIO_{i,g}$	Change in MSA level employment to population ratio between 2000 and 2010 for Black/ Hispanic workers (males and females), where i represents the MSA and g the vulnerable group.
	Primary variable of interest
$GSMI\_EFF_{i,h}$	Change in MSA SMI (jobs and housing units) between 2000 and 2010, where i represents the MSA, and h the housing unit type.

In our second specification, the labor market outcomes of the four vulnerable groups is represented by the following 'change' variables: $\Delta UNEMP\_RATE_{i,g,t}$ and $\Delta EMP\_POP\_RATIO_{i,g,t}$. The $\Delta UNEMP\_RATE_{i,g,t}$ represents the change in MSA level unemployment rate between 2000 and 2010 for the vulnerable group and $\Delta EMP\_POP\_RATIO_{i,g,t}$ represents the change in MSA level employment to population ratio between 2000 and 2010 for the vulnerable group.

$GSMI\_EFF_{i,h}$ is the primary variable of interest in this equation, which will enable us to evaluate the impact of the spatial mismatch between jobs and affordable housing units within an MSA on the labor market outcomes of vulnerable groups. $GSMI\_EFF_{i,h}$ represents the change in the level of GSMI between 2000 and 2010 in an MSA for a housing type. In equation 2a, a positive coefficient on the change in spatial mismatch between 2000 and 2000 (GSMI) can be viewed as evidence that the shift of job locations away from affordable housing locations negatively impacted labor market outcomes i.e. increases the unemployment rate of the vulnerable group. In equation 2b, a negative coefficient on the change in spatial mismatch between 2000 and 2000 (GSMI) can be viewed as evidence that the shift of job locations away from affordable housing locations negatively impacted labor market outcomes i.e. decreases the employment to population ratio of the vulnerable group.

iii. Model 3: Change in unemployment rate = f(Employment Shift, Housing Shift, and the changes in control variables)

$$\Delta UNEMP\_RATE_{i,g} = \alpha_i + \beta_1 * EMP\_EFF_{i,h} + \beta_2 * HSG\_EFF_{i,h} + \gamma * X_{i,t} + \varepsilon_{i,t} \tag{7}$$

$$\Delta EMP\_POP\_RATIO_{i,g} = \alpha_i + \beta_1 * EMP\_EFF_{i,h} + \beta_2 * HSG\_EFF_{i,h} + \gamma * X_{i,t} + \varepsilon_{i,t} \tag{8}$$

	Dependent variable
$\Delta UNEMP\_RATE_{i,g}$	Change in MSA level unemployment rate between 2000 and 2010 for Black/ Hispanic workers (males and females), where i represents the MSA and g the vulnerable group.
$\Delta EMP\_POP\_RATIO_{i,g}$	Change in MSA level employment to population ratio between 2000 and 2010 for Black/ Hispanic workers (males and females), where i represents the MSA and g the vulnerable group.
	Primary variables of interest
$EMP\_EFF_{i,h}$	Change in MSA SMI due to 2000–2010 employment shift (2000 SMI – mixed SMI), where i represents the MSA, and h the housing unit type.
$HSG\_EFF_{i,h}$	Change in MSA SMI due to 2000–2010 housing shift (mixed SMI – 2010 SMI), where i represents the MSA, and h the housing unit type.

In our third and final specification, the labor market outcomes of the four vulnerable groups and the array **X** of control variables remain unchanged from our previous specification.

In this model, the change in spatial mismatch between 2000 and 2000 (SMI or GSMI) is decomposed into the employment shift effect and the housing shift effect in order to assess the probable source of jobs – housing units spatial mismatch i.e. are jobs or housing is causing the shift in distribution. $EMP\_EFF_{i,h}$ represents MSA i's change in the SMI due to 2000–2010 employment shifts (2000 SMI – "Mixed SMI"), while the $HSG\_EFF_{i,h}$ represents MSA i's change in SMI due to 2000–2010 housing shifts ("Mixed SMI" – 2010 SMI). $EMP\_EFF_{i,h}$ and $HSG\_EFF_{i,h}$ are the primary variables of interest in this model, which are a decomposition of the $GSMI\_EFF_{i,h}$ used in model 2.

The employment shift effect is calculated as the difference between the 2000 SMI and the value of the SMI calculated using the 2010 employment distribution and the 2000 housing distribution (the "Mixed SMI"). A positive employment shift effect indicates that the changes in the distribution of metropolitan employment from 2000 to 2010 moved jobs farther away from affordable housing availability. Therefore, a positive coefficient on the employment shift effect variable implies that the unemployment rate for a vulnerable group tended to increase in metropolitan areas where the employment distribution shifted away from the affordable housing distribution. Therefore, a positive coefficient can be viewed as evidence that the shift of job locations away from affordable housing locations negatively impacted labor market outcomes.

The affordable housing shift effect is calculated by subtracting the "Mixed GSMI" from the 2010 GSMI. A negative affordable housing shift effect means that the changes in the metropolitan distribution of affordable housing between 2000 and 2010 lowered the GSMI and, therefore, brought the affordable housing distribution closer to the employment distribution, increasing job proximity. A positive coefficient, therefore, indicates that there was a tendency for unemployment rates to fall in metropolitan areas in which there was a negative affordable housing shift effect. This would suggest that the ability of workers to "follow" jobs to their new location helped to mitigate the impact of any employment shifts that reduced the job proximity of workers. Including both the employment shift effect and the affordable housing shift effect in the regression makes it possible to determine whether workers were able to offset any negative impact of employment shifts by moving to areas that were gaining jobs.

The focus of this section of the study is on the relationship between labor market outcomes and the distribution of jobs and affordable housing units within an MSA, it is important to control for other variables that determine performance of the labor market. It is important to note that in model 1, the control variables represent the level values of the variables for 2000 and 2010 in the regression equation. In the 2nd and 3rd model, the control variables represent the change in the values of the variables between 2000 and 2010. The array **X** contains a group of variables that are included as a control for the labor market conditions in the MSA.

	Control variables (X)
$LN\_POP_{i,t}$	Log (MSA population), where i represents the MSA and t the census year.
$POP\_GSMI_{i,r,t}$	MSA level GSMI between Black/Hispanic population and jobs, where i represents the MSA, r the race, and t the census year.
$POP\_SHR_{i,r,t}$	Black/Hispanic share of MSA population, where i represents the MSA, r the race, and t the census year.
$WHITE\_UNEMP_{i,t}$	MSA level unemployment rate for White males, where i represents the MSA and t the census year.
$POP\_HIGH\_ED_{i,r,t}$	Percentage of the MSA Black/Hispanic population over the age of 25 with at least a bachelor's degree, where i represents the MSA, r the race, and t the census year.
$PUBTRAN_{i,t}$	Percentage of MSA workers who use public transportation to commute to work, where i represents the MSA and t the census year.

$LN\_POP_{i,t}$ represents the log (MSA population), which is an indicator of the size of an MSA *i* and we include it in our regressions to control for the possibility that the size of the MSA affects labor market outcomes. is an important determinant of metropolitan economic performance and labor market outcomes. A higher $POP\_SHR_{i,r,t}$ is expected to negatively affect the labor market outcomes of the racial group we are evaluating, hence to control for this prospect we have included it as a control variable. $WHITE\_UNEMP_{i,t}$ represents the unemployment rate for White males in MSA *i*, which is an indicator of the health of the labor market in the MSA. $POP\_GSMI_{i,r,t}$ represents jobs and population GSMI for race *r* in MSA i and reflects the spatial mismatch between where jobs are located and where members of population group *r* reside. It measures the extent to which an MSA's geographical distribution of employment and race differ. This variable uses tract level employment and race data across 103 U.S. metropolitan statistical areas. This index is computed using the same methodology as the jobs – affordable housing index with the two groups being jobs –residents (Black/Hispanic). $POP\_GSMI_{i,r,t}$ is being included as a control in our model because Martin (2004) found that the increasing spatial separation of jobs and Black population, led to higher unemployment rates with the impact being most significant for younger workers.

$POP\_HIGH\_ED_{i,r,T}$ represents the percentage of MSA *i*'s population of race *r* greater than 25 years with at least a bachelor's degree, since the impact of spatial mismatch is expected to be lower for skilled workers, it is important to control for the education of the population of the race being evaluated. $PUBTRAN_{i,t}$ is the percentage of MSA workers that use public transportation to commute to work, where i represents the MSA and t the census year. It is included in the model to control for the level of public transportation across MSAs. Including the public transportation variable is important, because disadvantage workers may be more mobile in metropolitan areas with higher quality public transportation systems which could diminish the impact of housing spatial mismatch on labor market outcomes

In addition to the data obtained from the three primary data sources for 2000 and 2010 to compute the SMI and GSMI. The data for the variables in our estimating equation: unemployment data at the MSA level by race/gender, employment to population ratio data at the MSA level by race/gender, population share, percent of population using public transport and education attainment data is collected from the U.S. Census Bureau's American FactFinder database. Table 2.3 provides the basic summary statistics for the variables used in the regression section of this study. The statistics for these variables are calculated as weighted-averages using the respective population of each MSA as the weight. The unemployment rate for all four groups rose between 2000 and 2010. The employment to population ratio decreased for Black males and rose for the remaining three groups between 2000 and 2010. The population share of both Black and Hispanic residents rose between 2000 and 2010.

The percentage of Black and Hispanic population over the age of 25 years with a graduate degree or greater rose between 2000 and 2010. The unemployment rate for White males rose and the employment to population ratio fell between 2000 and 2010. The percentage of MSA population using public transportation was 3.82 percent in 2000. The average number of census tracts in the 103 MSAs with a population greater than 500,000 is 458.56.

The Employment – Race GSMI summary statistics reveals how the impact of employment shifts on GSMI values differs between typical worker in the MSA and Black/Hispanic workers. In 2000, the weighted average Black GSMI was 37.56 and Hispanic GSMI was 35.32, while the overall GSMI was 15.5. Therefore, the percentage of Black residents who would have to move in order to make their distribution across MSA census tracts identical to the distribution of employment was about 2.4 times higher than the percentage of all residents, while the percentage of Hispanic residents who would have to move in order to make their distribution across MSA census tracts identical to the distribution of employment was about 2.27 times higher than the percentage of all residents.

2.4.3 Results of regression models

To test the impact of the GSMI on the labor market performance of the four vulnerable groups of workers, we estimate OLS regression models (equations 3 through 8), which include various GSMI variables representing the disparity between jobs and housing units across the 103 MSAs. The results for the different specifications are presented in the following three sub sections.

i. Model 1 results:

In model 1, we have assessed the impact of jobs – housing shifts within an MSA on the labor market outcomes of Black and Hispanics. Tables 4 and 5 contain the summary results for the impact of the GSMI on the two labor market outcome variables for Blacks and Hispanics. For each race, the first column contains the results for the impact of jobs – housing shifts for the group containing affordable housing units at 80 percent of MFI, the second column contains the results for the group containing affordable housing units at 100 percent of MFI, the third column contains the results for the group containing affordable housing units at 120 percent of MFI, and the fourth and final column contains the results for the group containing all housing units within the MSA.

When evaluating the impact of the GSMI on the unemployment outcomes for Blacks and Hispanics, we find that the coefficient on the GSMI for *rental* units is positive when estimating its impact on the unemployment rate of Black males and females (except for Black females at 80 percent of MFI), the estimates for Hispanic males and females are statistically insignificant. In the case of the unemployment rate outcomes for *rental* units, the coefficient for Black males ranges from 0.000421 to 0.00234 and for Black females ranges from -0.000361 to 0.00102 (Table 2.4). The results for Black males and females are in line with our hypothesis and extant literature, which postulates that an increase in the spatial mismatch results in an increase in unemployment. Interestingly both Black males and females the results are significant at the upper end of affordability. Rental units at 80 percent of MFI have no impact on the unemployment level of Black males and rental units at 80 percent and 100 percent of MFI have no impact on the unemployment level of Black females

When evaluating the impact of the GSMI on the employment to population ratio for Blacks and Hispanics, we find that the coefficient on the GSMI for *rental* units is negative when estimating its impact on the unemployment rate of Black males, Black females and Hispanic females, the estimates for Hispanic males is statistically insignificant. In the case of the employment to population ratio for *rental* units, the coefficient for Blacks males ranges from -0.000229 to -0.00558, for Black females ranges from -0.00125 to -0.00361, and for Hispanic females ranges from -0.000107 to -0.00233 (Table 2.5). The results for Black males, Black females and Hispanic females are in line with our hypothesis and extant literature, which postulates that an increase in the spatial mismatch results in a decrease in the employment to population ratio. The employment to population ratio of Hispanic males is unaffected by an increase in the spatial mismatch.

On the other hand, we find that the coefficients on the GSMI for *owned* units are negative when estimating its impact on the unemployment rate and are positive when estimating its impact on the employment to population ratio for Blacks and Hispanics. In the case of the unemployment rate outcomes for *owned* units, the coefficient for Blacks males ranges from -0.0003 to -0.00135, for Black females ranges from 0.00006 to -0.00086, for Hispanic males ranges from -0.00024 to -0.000583, and for Hispanic females ranges from 0.00014 to -0.000554 (Table 2.4). These results are statistically significant across all three levels of affordability (except for 'All' units for all four vulnerable groups).

In the case of the employment to population ratio outcomes for *owned* units, the coefficient for Black males ranges from 0.00109 to 0.00163 and for Black females ranges from 0.000497 to 0.00107, while the coefficient for Hispanic males ranges from -0.000111 to 0.000622 and for Hispanic females ranges from -0.0001 to 0.00054 (Table 2.5). The statistical significance of the coefficients is mixed and the direction of the coefficients is not in line with our hypothesis.

We attribute these mixed results to endogeneity between home ownership and unemployment. Literature on home-ownership and unemployment rates was primarily driven by Oswald (1996, 1997, and 1999). The paper offers evidence that high rates of home-ownership result in high levels of unemployment. His hypothesis rests on the argument that home-owners are less mobile than renters. Therefore, an increase in the number of home-owners results in a decrease in the matching between job searchers and job opportunities, which results in higher unemployment of home-owners than that of renters. The two mechanisms at work are high transaction costs versus renters (more likely to refuse jobs that are farther away) and indirect effects (not move and maintain employment may require home-owners to commute longer distance than renters). Since the publication of Oswald's thesis, numerous studies have rejected his argument but have confirmed that home-ownership reduces mobility (Green and Hendershott (2001), Barrios and Rodriguez (2004)). This immobility of a home-owner causes endogeneity between home ownership and unemployment. Another source of endogeneity results from the fact that home-ownership and unemployment rates may be determined simultaneously.

Among the control variables, one measure has a consistent and statistically significant impact on the labor market outcomes of Blacks and Hispanics. The White unemployment rate/ employment to population ratio, which has a coefficient that is positive and statistically significant across all categories and vulnerable groups.

The coefficient for Black/Hispanic share of MSA population, is significant when estimating the unemployment rate of Blacks and the employment to population ratio of Blacks and Hispanics. An increase in the Black share of MSA population results in a decrease in the unemployment rate and an increase in the employment to population ratio. On the other hand an increase in the Hispanic share of MSA population does not impact the unemployment rate but results in an increase in the employment to population ratio.

An increase in the percentage of the MSA Black population over the age of 25 with at least a bachelor's degree, has a negative coefficient and is statistically significant with regard to the unemployment rate for Blacks and is statistically insignificant for Hispanics.

The dummy variable for differences between census year 2000 and 2010 is statistically significant across all specifications. Holding all other variables constant, the unemployment rate for the vulnerable groups reduced between 2000 and 2010 and the employment to population ratio increased across the same period.

ii. Model 2 results:

In model 2, we have assessed the impact of 2000 – 2010 change in the jobs – housing shifts within an MSA on the 2000 – 2010 change in labor market outcomes of Black and Hispanics (males and females). The first column contains the results for the impact of change in jobs – housing shifts for the group containing affordable housing units at 80 percent of MFI, the second column contains the results for the group containing affordable housing units at 100 percent of MFI, the third column contains the results for the group containing affordable housing units at 120 percent of MFI, and the fourth and final column contains the results for the group containing all housing units within the MSA.

In the case of the unemployment rate outcomes for *rental* units, the coefficient for Blacks males ranges from -0.00198 to -0.00309 and for Black females ranges from -0.00260 to -0.00387, while the coefficient for Hispanic males ranges from -0.00031 to -0.00075 and for Hispanic females ranges from -0.00134 to -0.00298 (Table 2.6). The direction of the coefficients is contrary to our hypothesis and the results are statistically insignificant (except for Hispanic females for 'All' units). In the case of the employment to population ratio outcomes for *rental* units, the coefficient for Black males ranges from 0.00076 to 0.00286 and for Black females ranges from 0.00154 to 0.00297, while the coefficient for Hispanic males ranges from -0.00027 to 0.00038 and for Hispanic females ranges from 0.00110 to 0.00384 (Table 2.7). Once again the direction of the coefficients is contrary to our hypothesis (except for Hispanic males at 120 percent of MFI and for 'All' units) and the results are statistically significant for Black males at 80 percent and 100 percent of MFI, Black females, and Hispanic females at 80 percent and 100 percent of MFI.

In the case of the unemployment rate outcomes for *owned* units, the coefficient for Black males ranges from -0.00206 to 0.00113 and for Black females ranges from -0.00140 to 0.00044, while the coefficient for Hispanic males ranges from -0.00079 to 0.00085 and for Hispanic females ranges from -0.00193 to 0.00123 (Table 2.6). These results are statistically significant (except for Black males at 80 percent of MFI and for 'All' units, Black females for 'All' units, Hispanic males, and Hispanic females for 'All' units) the direction of the coefficients is not in line with our hypothesis (except for 'All' units across all four vulnerable groups).

In the case of the employment to population ratio outcomes for *owned* units, the coefficient for Black males ranges from 0.00039 to 0.00114 and for Black females ranges from 0.00002 to 0.0016 (Table 2.7), while the coefficient for Hispanic males ranges from -0.00071 to 0.00063 and for Hispanic females ranges from -0.00289 to 0.00135 (Table 2.7). These results are statistically significant for Black males at 100 percent and 120 percent of MFI, Hispanic males for 'All' units and Hispanic females and the direction of the coefficients is not in line with our hypothesis (except for Hispanic males at 80 percent and 'All' units Hispanic females for 'All' units).

iii. Model 3 results:

Finally, we have assessed the impact of 2000 – 2010 change in employment and housing shifts within an MSA on the 2000 – 2010 change in labor market outcomes of Black and Hispanics (males and females). Tables 8 and 9 contain the summary results for the impact of the change in employment and housing shifts on the change in the two labor market outcome variables for Blacks and Hispanics. The first column contains the results for the impact of change in jobs – housing shifts for the group containing affordable housing units at 80 percent of MFI, the second column contains the results for the group containing affordable housing units at 100 percent of MFI, the third column contains the results for the group containing affordable housing units at 120 percent of MFI, and the fourth and final column contains the results for the group containing all housing units within the MSA.

In the case of the unemployment rate outcomes for *owned* units, the coefficient for employment shift GSMI for Black males ranges from -0.00265 to 0.00054 and the coefficient for housing shift GSMI ranges from -0.00213 to 0.00125, while the coefficient for employment shift SMI for Black females ranges from -0.0015 to -0.00087 and the coefficient for housing shift SMI ranges from -0.00136 to 0.0007 (Table 2.8). The coefficient for employment shift SMI for Hispanic males ranges from -0.00109 to 0.00077 and the coefficient for housing shift SMI ranges from -0.00084 to 0.00087, while the coefficient for employment shift SMI for Hispanic females ranges from -0.00453 to 0.00311 and the coefficient for housing shift SMI ranges from -0.00180 to 0.00066 (Table 2.8). These results are statistically insignificant (except for housing shift for Black males at 100 percent, 120 percent of MFI and housing shift for Black and Hispanic females at 80 percent, 100 percent, 120 percent of MFI).

In the case of the employment to population ratio outcomes for *owned* units, the coefficient for employment shift SMI for Black males ranges from -0.00146 to -0.00384 and the coefficient for housing shift SMI ranges from 0.000964 to 0.00142, while the coefficient for employment shift SMI for Black females ranges from -0.000014 to 0.000915 and the coefficient for housing shift SMI ranges from -0.000099 to 0.00117 (Table 2.9). These results are statistically insignificant (except for housing shift for Black males at 100 percent, and 120 percent of MFI). The coefficient for employment shift SMI for

Hispanic males ranges from -0.0101 to -0.0084 and the coefficient for housing shift SMI ranges from -0.00156 to 0.00074, while the coefficient for employment shift SMI for Hispanic females ranges from -0.00724 to -0.00573 and the coefficient for housing shift SMI ranges from -0.00175 to 0.00167 (Table 2.9). These results are statistically significant for Hispanic males and females (except for housing shift for Hispanic males, employment shift for Hispanic females at 80 percent, of MFI and housing shift for Hispanic females at 'All' units).

In the case of the unemployment rate outcomes for *rental* units, the coefficient for employment shift GSMI for Black males ranges from -0.00154 to -0.0009 and the coefficient for housing shift GSMI ranges from -0.00367 to -0.0023, while the coefficient for employment shift GSMI for Black females ranges from -0.00268 to -0.00192 and the coefficient for housing shift GSMI ranges from -0.00432 to -0.00267 (Table 2.8). These results are statistically significant (except for employment shift for Black males and females and housing shift for Black males at 'All' units).

The coefficient for employment shift GSMI for Hispanic males ranges from -0.00082 to -0.00009 and the coefficient for housing shift SMI ranges from -0.00083 to -0.00002, while the coefficient for employment shift SMI for Hispanic females ranges from -0.00019 to 0.00027 and the coefficient for housing shift SMI ranges from -0.00387 to -0.00285 (Table 2.8). These results are statistically insignificant (except for employment shift for Hispanic females at 80 percent, 100 percent, and 120 percent of MFI).

In the case of the employment to population ratio outcomes for *rental* units, the coefficient for employment shift GSMI for Black males ranges from 0.00221 to 0.00236 and the coefficient for housing shift SMI ranges from 0.00056 to 0.0031, while the coefficient for employment shift SMI for Black females ranges from 0.0000 to 0.00166 and the coefficient for housing shift SMI ranges from 0.00153 to 0.00366 (Table 2.9). These results are statistically insignificant for both the employment shift and housing shift for Black males. The results are insignificant for employment shift and significant for the housing shift for Black females.

The coefficient for employment shift SMI for Hispanic males ranges from 0.00161 to 0.00530 and the coefficient for housing shift SMI ranges from -0.00039 to -0.00021, while the coefficient for employment shift SMI for Hispanic females ranges from 0.00195 to 0.00534 and the coefficient for housing shift SMI ranges from 0.00025 to 0.00389 (Table 2.9). These results are statistically insignificant for Hispanic males and females (except for employment shift for Hispanic females at 80 percent of MFI and Housing shift for Hispanic females at 80 percent, and 100 percent of MFI).

2.4.4 Summary:

The results of the regression analysis are mixed. When evaluating the impact of GSMI shift on the levels of unemployment rate/ employment to population ratio in 2000 and 2010 (model 1), Black males and females are positively impacted by an increase in the GSMI between rental units and jobs, but the impact on Hispanic males and females is statistically insignificant. The employment to population ratio for Black males, Black females, and Hispanic females decreases when the GSMI between rental units and jobs increases. The employment to population ratio for Hispanic males increases but the results are statistically insignificant. For owned units, the unemployment rate for all four groups decreases when the GSMI increases and the employment to population ratio for all four groups increases when GSMI increases. The results for owned units are contrary to our hypothesis and we attribute them to endogeneity between home ownership and unemployment, which is caused due to simultaneity and immobility due to homeownership.

When evaluating the impact of the total GSMI shift on the change in unemployment rate/ employment to population ratio between 2000 and 2010 (model 2), for rental units, the change in the unemployment rate for Black (males and females) and Hispanics females decreases when the change in GSMI increases and the change in the employment to population ratio for Black (males and females) and Hispanic females increase when the change in GSMI increases. The impact of total GSMI shift on Hispanic males is insignificant.

For ownership units, the change in the unemployment rate for all four groups Black and Hispanics (males and females) decreases when the change in GSMI increases and the change in the employment to population ratio for Black (males and females) and Hispanic females increases when the change in GSMI increases. Once again, we believe that these results inconsistent with our hypothesis are a result of endogeneity and simultaneity between home ownership and employment.

This leads us to conclude that an increase in GSMI between affordable housing and jobs does not negatively impact the labor market outcomes of the four vulnerable groups we have studied in this paper. Instead, over the 10 year period, the change in labor market outcomes are inversely related to GSMI. These results do not support the hypothesis that a spatial match between affordable housing and jobs negatively affects the labor market outcomes of the four vulnerable groups between 2000 and 2010 across 103 MSAs in the US.

When evaluating the impact of the employment shift and the housing shift on the change in labor market outcomes between 2000 and 2010 (model 3), for rental units, we find that when the change in employment shift GSMI does not impact the change in the unemployment rate for Black and Hispanics (males and females) decreases and when the change in housing shift GSMI increases, the change in the unemployment rate for Black (males and females) and Hispanic females decreases.

When evaluating owned units, the change in the unemployment rate for Black and Hispanics (males and females) is not impacted by the change in employment shift GSMI and the change in housing shift GSMI negatively impacts the unemployment rate of Black males (for units affordable at 100 percent and 120 percent of area median income) and Hispanic females.

For rental units, the change in employment shift GSMI does not have an impact on the employment to population ratio of the four vulnerable groups except for Hispanic females for units affordable at 80 percent of area median income). The change in housing shift GSMI positively impacts the employment to population ratio of Black females and Hispanic females for units affordable at 80 percent and 100 percent of area median income. When shifting our attention to owned units, we find that a change in employment shift GSMI negatively impacts the employment to population ratio of Black females and Hispanic males and females. Black males are not impacted by the change in the employment shift GSMI as the results are statistically insignificant. The change in the housing shift GSMI positively impacts the employment to population ratio of Black males (for units affordable at 100 percent and 120 percent of area median income) and Hispanic females (for units affordable at 100 percent and 120 percent of area median income).

When evaluating the impact of the employment shift and the housing shift, the results do not support the hypothesis that an increase in the GSMI negatively affects the labor market outcomes of the four vulnerable groups between 2000 and 2010 across 103 MSAs in the US.

2.5 Conclusion

The movement of jobs shifting away from the city center to the suburbs in nearly all U.S. metropolitan area began in the 1950s and has continued as recently as the 1990s and 2000s. In 1968, John Kain, while studying this occurrence, concluded that racial discrimination in housing markets affected the spatial distribution of Black employment and reduced Black employment in metropolitan areas. Over the years a steady stream of literature has analyzed the underlying causes of spatial mismatch in metropolitan areas and suggested ways to improve the poor labor market outcomes for the affected groups. This paper studied one possible explanation for the persistence of spatial mismatch across US metropolitan areas – mismatch between the distribution of affordable housing and jobs between 2000 and 2010.

This paper expands upon existing spatial mismatch studies in three ways. First, it examines whether there is a relationship between spatial mismatch and the availability of affordable housing in US metropolitan areas. Second, it constructs the spatial mismatch indices using tract-level data, which allows us to measure employment and housing shifts uniformly across neighborhoods within the metropolitan area. A census tract is significantly smaller in size as compared to a county, which allows us to define distinct neighborhoods and the spatial interaction between the neighborhoods within an MSA. Lastly, it uses a distance-weighted index when computing the general spatial mismatch index. This is helpful because a distance-weighted – General Spatial Mismatch Index (GSMI) – captures the spatial interaction between census tracts within an MSA and the spatial interaction between census tracts decreases as distance increases.

Using tract-level employment and housing data for 103 U.S. metropolitan statistical areas we create a traditional spatial mismatch index (SMI) and a distance weighted spatial mismatch index (GSMI) to analyze whether jobs are moving to areas where affordable housing is less available. We find that the results of the SMI and GSMI across three categories of affordability, for both rental and ownership units, are diametrically opposite for certain categories, i.e., spatial mismatch has decreased under SMI and increased under GSMI. Interestingly, incorporating a distance decay measure reveals the true picture with regard to the level of mismatch between jobs and the distribution of affordable housing within an MSA.

We then test whether there is a relationship between the extent to which jobs move to areas with fewer affordable housing units and the labor market outcomes of potentially disadvantaged workers. The paper finds that, on average, using levels in 2000 and 2010, the labor market outcomes of Black males and females were impacted by an increase in the spatial mismatch between affordable rental housing and jobs, while Hispanic males and females were unaffected.

We then evaluated the impact of the change in GSMI between 2000 and 2010 on the change in the labor outcomes between 2000 and 2010. The labor market outcomes improved with an increase in the spatial mismatch between affordable housing and jobs for Black (males and females) and Hispanic females. Hispanic males were not affected by the change in the spatial mismatch between affordable housing and jobs. Our results are inconsistent with previous results and provide no evidence that spatial mismatch between affordable housing and jobs contributes to the persistence of poor labor market outcomes of Blacks and Hispanics.

Although, this paper did not find evidence that spatial mismatch between affordable housing and jobs within US metropolitan areas plays a role in the labor market outcomes of the four vulnerable groups – Black males, Black females, Hispanic males and Hispanic females, this result is significant. It confirms that a potential source for the persistence of poor labor market outcomes as a result of jobs moving away from affordable housing or the distribution of affordable housing shifting away from the location of jobs is not a cause.

2.6 References:

Blumenberg, E. (2004). En-gendering Effective Planning: Spatial Mismatch, Low-Income Women, and Transportation Policy. *Journal- American Planning Association*, (3), 269.

Blumenberg, E., & Pierce, G. (2014). Multimodal travel and the poor: evidence from the 2009 National Household Travel Survey. *Transportation Letters: The International Journal of Transportation Research*, 6(1), 36.

Carlson, V. L., & Theodore, N. (1997). Employment Availability for Entry-Level Workers: An Examination of the Spatial-Mismatch Hypothesis in Chicago. *Urban Geography*, (3), 228.

Cervero, R., & Rood, T. (1999). Tracking accessibility: employment and housing opportunities in the San Francisco Bay Area. *Environment & Planning A*, *31*(7), 1259.

Chapple, K. (2006). Overcoming Mismatch. *Journal of the American Planning Association*, *72*(3), 322–336.

Dodson, J. (2005) Is there a spatial mismatch between housing affordability and employment opportunity in Melbourne? *AHURI Final Report No. 80*, Australian Housing and Urban Research Institute Limited, Melbourne.

Ellen, I. G., & Turner, M. A. (1997). Does Neighborhood Matter? Assessing Recent Evidence. *Housing Policy Debate -Washington-*, (4), 833.

Ellwood, D. T. (1986). The Spatial Mismatch Hypothesis: Are There Teenage Jobs Missing in the Ghetto? In R. B. Freeman & H. J. Holzer (Eds.), *The black youth employment crisis* (pp. 147–185). National Bureau of Economic Research Project Report series.

Fan, Y., and N. Tilahun (2014). Transit and job accessibility: an empirical study of access to competitive clusters and regional growth strategies for enhancing transit accessibility. *Transport Policy*, *33*, 17–25.

Fan, Y. (2012). The Planners' War against Spatial Mismatch: Lessons Learned and Ways Forward. *Journal of Planning Literature*, (2), 153.

Fan, Y., Andrew Guthrie and David Levinson, 2010. "Impact of Light Rail Implementation on Labor Market Accessibility: A Transportation Equity Perspective," Working Papers 000079, University of Minnesota: Nexus Research Group.

Fieldhouse, Edward A. (1999). Ethnic Minority Unemployment and Spatial Mismatch: The Case of London. *Urban Studies*, (9), 1569.

Harry J. Holzer. (1991). The Spatial Mismatch Hypothesis: What Has the Evidence Shown? *Urban Studies*, (1), 105.

Donald S. Houston. (2005). Methods to Test the Spatial Mismatch Hypothesis. *Economic Geography*, (4), 407.

Donald Houston. (2005). Employability, Skills Mismatch and Spatial Mismatch in Metropolitan Labour Markets. *Urban Studies*, (2), 221.

Ihlanfeldt, K., & Sjoquist, D. (1998). The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform. *Housing Policy Debate*, *9*(4), 849.

Ihlanfeldt, K. (1994). The Spatial Mismatch between Jobs and Residential Locations within Urban Areas. *Cityscape: A Journal of Policy Development and Research*, *1*(1), 219–244.

Immergluck, D. (1998). *Neighborhood Jobs, Race, and Skills : Urban Unemployment and Commuting*. New York : Garland Pub., 1998.

Kain, John F. (1968). Housing Segregation, Negro Employment, and Metropolitan Decentralization. *The Quarterly Journal of Economics*, (2), 175.

Kawabata, M., & Qing Shen. (2006). Job accessibility as an indicator of auto-oriented urban structure: a comparison of Boston and Los Angeles with Tokyo.

Liu, Z., and Wang, M. (2011). Job Accessibility and Its Impacts on Commuting Time of Urban Residents in Beijing: From a Spatial Mismatch Perspective. *Acta Geographica Sinica -Chinese Edition-*, (4), 467.

Martin, R. W. (2004). Can Black workers escape spatial mismatch? Employment shifts, population shifts, and Black unemployment in American cities. *Journal of Urban Economics*, (1), 179.

Moss, P., and Tilly, C. (1993). A Turn for the Worse: Why Black Men's Labour Market Fortunes Have Declined In the United States. *Sage Race Relations Abstracts*, *18*(1), 5.

Musterd, S., and de Winter, M. (1998). Conditions for Spatial Segregation: Some European Perspectives. *International Journal of Urban & Regional Research*, *22*(4), 665–673.

Painter, Gary, Cathy Yang Liu, and Duan Zhuang. (2007). Immigrants and the Spatial Mismatch Hypothesis: Employment Outcomes among Immigrant Youth in Los Angeles. *Urban Studies*, (13), 2627.

Stoll, M. A. (1999). Spatial Job Search, Spatial Mismatch, and the Employment and Wages of Racial and Ethnic Groups in Los Angeles. *Journal of Urban Economics*, (1), 129.

Taylor, Brian D. and Paul M. Ong. (1995). Spatial Mismatch or Automobile Mismatch? An Examination of Race, Residence and Commuting in US Metropolitan Areas. *Urban Studies*, (9), 1453.

Wheeler, Laura A. (1990). A Review of the Spatial Mismatch Hypothesis: Its Impact on the Current Plight of the Central City in the United States. Occasional Paper No. 137. Syracuse University, Maxwell School of Citizenship and Public Affairs, Metropolitan Studies Program.

Table 2.1: Weighted average Spatial Mismatch Indices (SMIs) and the General Spatial Mismatch Indices (GSMIs) for rental units across the 103 largest MSAs in the US.
The first set of columns represents the mean and SD for the year 2000 and the second the mean and SD for the year 2010. The final set of columns present the difference in the means and SDs between 2010 and 2000. The top half of the table presents the results of the SMI and the bottom half the GSMI. The change in the SMI value is decomposed into the Employment Shift (difference in the "mixed SMI/GSMI" and the) and the Housing Shift (difference in the "mixed SMI/GSMI" and the 2010 SMI/GSMI). The respective groups are further divided into SMI/GSMI for all units, units affordable at 80 percent, 100 percent and 120 percent of median household income. The mean and standard deviation for each of the variables were calculated by weighting each of the 103 MSAs in the sample by their 2010 population.

	2000		**2010**		**Change**	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
SMI Rental Units:						
All Units	45.27	6.16	43.75	5.39	-1.52	1.68
Mixed SMI			45.43	5.35		
Employment Shift					*0.16*	*1.65*
Housing Shift					*-1.68*	*1.47*
Units @ 80% MFI	44.43	5.88	42.30	4.55	-2.13	2.03
Mixed SMI			44.70	5.12		
Employment Shift					*0.27*	*1.57*
Housing Shift					*-2.40*	*1.49*
Units @ 100% MFI	44.98	6.05	43.07	4.91	-1.92	1.84
Mixed SMI			45.20	5.26		
Employment Shift					*0.21*	*1.62*
Housing Shift					*-2.13*	*1.41*
Units @ 120% MFI	45.17	6.10	43.44	5.12	-1.73	1.72
Mixed SMI			45.35	5.31		
Employment Shift					*0.18*	*1.63*
Housing Shift					*-1.91*	*1.37*
GSMI Rental Units:						
All Units	12.73	3.60	12.29	3.10	-0.44	1.11
Mixed SMI			12.79	3.14		
Employment Shift					*0.06*	*0.96*
Housing Shift					*-0.51*	*0.93*
Units @ 80% MFI	13.12	3.59	14.01	3.80	0.88	2.35
Mixed SMI			13.25	3.15		
Employment Shift					*0.12*	*0.90*
Housing Shift					*0.76*	*2.23*
Units @ 100% MFI	12.82	3.73	12.72	3.22	-0.10	1.38
Mixed SMI			12.91	3.27		
Employment Shift					*0.09*	*0.95*
Housing Shift					*-0.19*	*1.19*
Units @ 120% MFI	12.77	3.72	12.43	3.22	-0.34	1.17
Mixed SMI			12.85	3.25		
Employment Shift					*0.08*	*0.96*
Housing Shift					*-0.41*	*0.98*

Table 2.2: Weighted average Spatial Mismatch Indices (SMIs) and the General Spatial Mismatch Indices (GSMIs) for ownership units across the 103 largest MSAs in the US.
The first set of columns represents the mean and SD for the year 2000 and the second the mean and SD for the year 2010. The final set of columns present the difference in the means and SDs between 2010 and 2000. The top half of the table presents the results of the SMI and the bottom half the GSMI. The change in the SMI value is decomposed into the Employment Shift (difference in the "mixed SMI/GSMI" and the 2000 SMI/GSMI) and the Housing Shift (difference in the "mixed SMI/GSMI" and the 2010 SMI/GSMI). The respective groups are further divided into SMI/GSMI for all units, units affordable at 80 percent, 100 percent and 120 percent of median household income. The mean and standard deviation for each of the variables were calculated by weighting each of the 103 MSAs in the sample by their 2010 population.

	2000		**2010**		**Change**	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
SMI Ownership Units:						
All Units	49.94	5.49	47.17	4.21	-2.77	2.66
Mixed SMI			48.53	5.30		
Employment Shift					*-1.41*	*1.53*
Housing Shift					*-1.36*	*2.87*
Units @ 80% MFI	43.64	3.54	42.27	2.98	-1.37	1.95
Mixed SMI			43.20	3.36		
Employment Shift					*-0.44*	*1.12*
Housing Shift					*-0.93*	*1.95*
Units @ 100% MFI	44.71	4.03	42.65	3.19	-2.06	2.25
Mixed SMI			43.93	3.77		
Employ1ment Shift					*-0.78*	*1.26*
Housing Shift					*-1.28*	*2.31*
Units @ 120% MFI	45.87	4.24	43.22	3.39	-2.64	2.34
Mixed SMI			44.88	4.00		
Employment Shift					*-0.99*	*1.32*
Housing Shift					*-1.65*	*2.51*
GSMI Ownership Units:						
All Units	21.81	6.99	19.17	3.73	-2.63	4.69
Mixed SMI			21.02	7.45		
Employment Shift					*-0.78*	*1.39*
Housing Shift					*-1.85*	*5.32*
Units @ 80% MFI	26.52	8.81	28.09	9.53	1.57	2.97
Mixed SMI			26.37	8.92		
Employment Shift					*-0.16*	*0.72*
Housing Shift					*1.72*	*2.97*
Units @ 100% MFI	23.31	7.22	25.08	8.52	1.76	3.48
Mixed SMI			22.96	7.44		
Employment Shift					*-0.35*	*0.99*
Housing Shift					*2.12*	*3.47*
Units @ 120% MFI	21.97	6.35	23.21	7.38	1.24	4.42
Mixed SMI			21.49	6.65		
Employment Shift					*-0.48*	*1.16*
Housing Shift					*1.72*	*4.45*

Table 2.3: Summary Statistics across the 103 largest MSAs in the US.
The mean and standard deviation for each of the variables were calculated by weighting each of the 103 metropolitan areas in the sample by their respective populations.

	2000		**2010**		**Change**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Unemployment rate:						
Black males	12.040	2.633	15.366	3.391	3.326	2.809
Black females	10.469	1.830	12.461	2.513	1.992	2.388
Hispanic males	8.085	2.452	8.972	2.022	0.888	2.123
Hispanic females	10.383	2.798	10.386	2.088	0.004	2.353
Employment to Population Ratio:						
Black males	55.411	5.139	55.264	5.256	-0.147	3.045
Black females	54.661	3.964	56.431	4.015	1.770	2.437
Hispanic males	63.762	4.976	70.770	5.418	7.008	2.843
Hispanic females	47.676	4.374	52.982	3.298	5.307	2.679
Ln (MSA population)	14.890	1.065	14.980	1.042	0.090	0.081
Population Share:						
Black	18.989	8.966	19.379	9.191	0.389	1.256
Hispanic	26.686	16.815	30.681	17.105	3.995	1.837
Percentage of population over 25 with a Graduate degree or greater:						
Blacks	1.201	2.235	1.528	2.796	0.327	0.669
Hispanics	1.096	1.606	1.456	2.080	0.360	0.563
Unemployment rate for White males	4.008	0.880	6.731	1.346	2.723	1.185
Employment to Population Ratio for White males	70.699	4.638	67.450	4.122	-3.249	1.733
Percentage of commuters using public transportation	3.816	4.310	–	–	–	–
Number of census tracts in MSA	458.56	609.69	–	–	–	–

Table 2.4: Summary results for the impact of the general spatial mismatch index (GSMI) on the unemployment rate of Blacks and Hispanics.
*This table reports the pooled OLS regression coefficients for the variables of interest with their corresponding t-stats below them in parentheses for three different affordable housing levels and the four vulnerable groups (Black males, Black females, Hispanic males, and Hispanic females). These specifications are from left to right; units available at 80 percent of median household income, 100 percent of median household income, 120 percent of median household income, and all rental units in the metropolitan area. Statistical significance is designated by *** at the 1% level, ** at the 5% level, and * at the 10% level. The complete set of results for each individual OLS regression is available upon request.*

		Blacks				**Hispanics**			
Group	**Parameter**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**
Males	Rental Units - Jobs GSMI	0.000421	0.00193***	0.00212***	0.00234***	-0.000264	-0.000001	-0.0000985	-0.000185
		(0.65)	(2.98)	(3.39)	(3.79)	(-0.66)	(-0.00)	(-0.21)	(-0.39)
	Owned Units - Jobs GSMI	-0.00104***	-0.00121***	-0.00135***	-0.00030	-0.000516***	-0.000583***	-0.000571***	-0.00024
		(-3.65)	(-3.78)	(-3.92)	(-0.85)	(-3.25)	(-3.28)	(-2.98)	(-0.92)
Females	Rental Units - Jobs GSMI	-0.000361	0.00064	0.000843*	0.00102**	-0.000131	0.000372	0.000406	0.000235
		(-0.75)	(1.30)	(1.77)	(2.17)	(-0.26)	(0.62)	(0.67)	(0.39)
	Owned Units - Jobs GSMI	-0.000805***	-0.000825***	-0.000862***	0.0000638	-0.000518**	-0.000552**	-0.000554**	0.000148
		(-3.80)	(-3.44)	(-3.32)	(0.24)	(-2.51)	(-2.39)	(-2.23)	(0.44)

Table 2.5: Summary results for the impact of the general spatial mismatch index on the employment to population ratio of Blacks and Hispanics. *This table reports the pooled OLS regression coefficients for the variables of interest with their corresponding t-stats below them in parentheses for three different affordable housing levels and the four vulnerable groups (Black males, Black females, Hispanic males, and Hispanic females). These specifications are from left to right; units available at 80 percent of median household income, 100 percent of median household income, 120 percent of median household income, and all rental units in the metropolitan area. Statistical significance is designated by *** at the 1% level, ** at the 5% level, and * at the 10% level. The complete set of results for each individual OLS regression is available upon request.*

		Blacks				**Hispanics**			
Group	**Parameter**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**
Males	Rental Units - Jobs GSMI	-0.00229*	-0.00488***	-0.00526***	-0.00558***	0.00085	0.000980	0.000901	0.000811
		(-1.76)	(-3.86)	(-4.32)	(-4.61)	(0.82)	(0.81)	(0.73)	(0.66)
	Owned Units - Jobs GSMI	0.00109*	0.00097	0.00136*	0.00163**	-0.000111	0.000211	0.000557	0.000622
		(1.84)	(1.46)	(1.91)	(2.30)	(-0.27)	(0.44)	(1.07)	(0.95)
Females	Rental Units - Jobs GSMI	-0.00125	-0.00309***	-0.00339***	-0.00361***	-0.000107	-0.00172*	-0.00233**	-0.00231**
		(-1.23)	(-3.10)	(-3.52)	(-3.77)	(-0.13)	(-1.79)	(-2.41)	(-2.38)
	Owned Units - Jobs GSMI	0.000497	0.000237	0.000413	0.00107*	0.000540*	0.000465	0.000395	-0.00100*
		(1.07)	(0.46)	(0.74)	(1.93)	(1.67)	(1.21)	(0.95)	(-1.93)

Table 2.6: Summary results for the impact of the change in the general spatial mismatch index on the unemployment rate of Blacks and Hispanics. *This table reports the OLS regression coefficients for the variables of interest with their corresponding t-stats below them in parentheses for three different affordable housing levels and the four vulnerable groups (Black males, Black females, Hispanic males, and Hispanic females). These specifications are from left to right; units available at 80 percent of median household income, 100 percent of median household income, 120 percent of median household income, and all rental units in the metropolitan area. Statistical significance is designated by *** at the 1% level, ** at the 5% level, and * at the 10% level. The complete set of results for each individual OLS regression is available upon request.*

		Blacks				**Hispanics**			
Group	**Parameter**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**
Males	Total Shift GSMI (Rental Units)	-0.00227***	-0.00309**	-0.00245	-0.00198	-0.00075	-0.00053	-0.00052	-0.00031
		(-3.18)	(-2.34)	(-1.60)	(-1.22)	(-1.46)	(-0.57)	(-0.42)	(-0.23)
	Total Shift GSMI (Ownership Units)	-0.000958	-0.00208***	-0.00144***	0.00113	-0.000841	-0.000366	-0.000127	0.000847
		(-1.61)	(-3.66)	(-3.15)	(1.51)	(-1.63)	(-1.05)	(-0.48)	(1.04)
Females	Total Shift GSMI (Rental Units)	-0.00260***	-0.00387***	-0.00335**	-0.00285**	-0.00271***	-0.00298***	-0.00227	-0.00134
		(-4.25)	(-3.40)	(-2.50)	(-2.00)	(-4.77)	(-2.72)	(-1.50)	(-0.79)
	Total Shift SMI (Ownership Units)	-0.00138***	-0.00145***	-0.000938**	0.000437	-0.00192***	-0.00123***	-0.000682**	0.00123
		(-2.66)	(-2.80)	(-2.25)	(0.65)	(-3.14)	(-2.99)	(-2.16)	(1.24)

Table 2.7: Summary results for the impact of the change in the general spatial mismatch index on the employment to population ratio of Blacks and Hispanics.

*This table reports the OLS regression coefficients for the variables of interest with their corresponding t-stats below them in parentheses for three different affordable housing levels and the four vulnerable groups (Black males, Black females, Hispanic males, and Hispanic females). These specifications are from left to right; units available at 80 percent of median household income, 100 percent of median household income, 120 percent of median household income, and all rental units in the metropolitan area. Statistical significance is designated by *** at the 1% level, ** at the 5% level, and * at the 10% level. The complete set of results for each individual OLS regression is available upon request.*

		Blacks				Hispanics			
Group	**Parameter**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**
Males	Total Shift GSMI (Rental Units)	0.00076	0.00222	0.00280*	0.00286*	0.00038	0.00025	-0.00027	-0.00116
		(0.91)	(1.55)	(1.77)	(1.72)	(0.35)	(0.13)	(-0.11)	(-0.44)
	Total Shift GSMI (Ownership Units)	0.000439	0.00114*	0.00112**	0.000627	-0.00071	0.00046	0.00063	-0.00282*
		(0.63)	(1.78)	(2.32)	(0.81)	(-0.72)	(0.70)	(1.33)	(-1.94)
Females	Total Shift GSMI (Rental Units)	0.00154*	0.00272**	0.00294*	0.00297*	0.00302***	0.00384***	0.00259	0.00110
		(1.97)	(2.03)	(1.97)	(1.90)	(3.76)	(2.68)	(1.36)	(0.52)
	Total Shift SMI (Ownership Units)	0.00106	0.000473	0.000226	0.0000182	0.00131*	0.00135***	0.00106***	-0.00289**
		(1.63)	(0.78)	(0.48)	(0.02)	(1.68)	(2.67)	(2.86)	(-2.49)

Table 2.8: Summary results for the impact of the decomposed general spatial mismatch index on the unemployment rate of Blacks and Hispanics. *This table reports the OLS regression coefficients for the variables of interest with their corresponding t-stats below them in parentheses for three different affordable housing levels and the four vulnerable groups (Black males, Black females, Hispanic males, and Hispanic females). These specifications are from left to right; units available at 80 percent of median household income, 100 percent of median household income, 120 percent of median household income, and all rental units in the metropolitan area. Statistical significance is designated by *** at the 1% level, ** at the 5% level, and * at the 10% level. The complete set of results for each individual OLS regression is available upon request.*

		Blacks				**Hispanics**			
Group	**Parameter**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**
Males	Employment Shift GSMI (Rental Units)	-0.00141 (-0.64)	-0.00154 (-0.68)	-0.00116 (-0.50)	-0.000909 (-0.38)	-0.00009 (-0.04)	-0.00076 (-0.41)	-0.00082 (-0.45)	-0.00063 (-0.34)
	Housing Shift GSMI (Rental Units)	-0.00235*** (-3.15)	-0.00367*** (-2.47)	-0.00315* (-1.75)	-0.00258 (-1.36)	-0.00083 (-1.46)	-0.00045 (-0.43)	-0.00031 (-0.20)	-0.00002 (-0.01)
	Employment Shift GSMI (Ownership Units)	-0.00265 (-0.94)	-0.00157 (-0.72)	-0.00126 (-0.62)	0.000535 (0.31)	-0.000895 (-0.35)	-0.00109 (-0.49)	-0.000459 (-0.22)	0.000770 (0.50)
	Housing Shift GSMI (Ownership Units)	-0.000797 (-1.22)	-0.00213*** (-3.50)	-0.00145*** (-3.08)	0.00125 (1.54)	-0.000838 (-1.57)	-0.000342 (-0.95)	-0.000120 (-0.45)	0.000870 (0.96)
Females	Employment Shift GSMI (Rental Units)	-0.00192 (-1.01)	-0.00268 (-1.38)	-0.00267 (-1.32)	-0.00259 (-1.25)	0.00062 (0.27)	-0.00019 (-0.09)	-0.00016 (-0.07)	0.00027 (0.12)
	Housing Shift GSMI (Rental Units)	-0.00267*** (-4.16)	-0.00432*** (-3.36)	-0.00371** (-2.36)	-0.00300* (-1.80)	-0.00311*** (-5.00)	-0.00387*** (-3.13)	-0.00375** (-2.00)	-0.00285 (-1.29)
	Employment Shift GSMI (Ownership Units)	-0.00155 (-0.63)	-0.00212 (-1.06)	-0.00208 (-1.14)	-0.000869 (-0.57)	-0.00453 (-1.49)	-0.000977 (-0.37)	0.000664 (0.27)	0.00311 (1.66)
	Housing Shift GSMI (Ownership Units)	-0.00136** (-2.40)	-0.00138** (-2.49)	-0.000878** (-2.05)	0.000697 (0.96)	-0.00180*** (-2.86)	-0.00124*** (-2.93)	-0.000709** (-2.21)	0.000662 (0.60)

Table 2.9: Summary results for the impact of the decomposed general spatial mismatch index on the employment to population ratio of Blacks and Hispanics.

*This table reports the OLS regression coefficients for the variables of interest with their corresponding t-stats below them in parentheses for three different affordable housing levels and the four vulnerable groups (Black males, Black females, Hispanic males, and Hispanic females). These specifications are from left to right; units available at 80 percent of median household income, 100 percent of median household income, 120 percent of median household income, and all rental units in the metropolitan area. Statistical significance is designated by *** at the 1% level, ** at the 5% level, and * at the 10% level. The complete set of results for each individual OLS regression is available upon request.*

		Blacks				**Hispanics**			
Group	**Parameter**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**	**Units @ 80% MFI**	**Units @ 100% MFI**	**Units @ 120% MFI**	**All Units**
Males	Employment Shift GSMI	0.00236	0.00221	0.00227	0.00233	0.00530	0.00278	0.00222	0.00161
	(Rental Units)	(0.99)	(0.94)	(0.96)	(0.97)	(1.43)	(0.83)	(0.66)	(0.48)
	Housing Shift GSMI	0.00056	0.00223	0.00310	0.00316	-0.00021	-0.00063	-0.00216	-0.00390
	(Rental Units)	(0.63)	(1.36)	(1.65)	(1.62)	(-0.18)	(-0.30)	(-0.73)	(-1.15)
	Employment Shift GSMI	-0.00384	-0.00213	-0.00158	-0.00146	-0.0125***	-0.0101***	-0.00840**	-0.00755***
	(Ownership Units)	(-1.38)	(-0.93)	(-0.76)	(-0.84)	(-2.92)	(-2.63)	(-2.42)	(-2.84)
	Housing Shift GSMI	0.000881	0.00142**	0.00123**	0.000964	-0.00011	0.00073	0.00074	-0.00156
	(Ownership Units)	(1.18)	(2.15)	(2.52)	(1.19)	(-0.11)	(1.16)	(1.62)	(-1.01)
Females	Employment Shift GSMI	0.00000	0.00161	0.00166	0.00180	0.00534*	0.00370	0.00289	0.00195
	(Rental Units)	(0.71)	(0.73)	(0.75)	(0.79)	(1.90)	(1.41)	(1.07)	(0.71)
	Housing Shift GSMI	0.00153*	0.00318**	0.00366**	0.00365**	0.00275***	0.00389**	0.00236	0.00025
	(Rental Units)	(1.85)	(2.08)	(2.08)	(1.99)	(3.18)	(2.39)	(0.99)	(0.09)
	Employment Shift GSMI	0.0000142	0.000776	0.000915	0.000743	-0.00573	-0.00724**	-0.00615**	-0.00717***
	(Ownership Units)	(0.01)	(0.35)	(0.45)	(0.45)	(-1.64)	(-2.41)	(-2.27)	(-3.41)
	Housing Shift GSMI	0.00117	0.000446	0.000198	-0.0000992	0.00167**	0.00157***	0.00115***	-0.00175
	(Ownership Units)	(1.66)	(0.70)	(0.42)	(-0.13)	(2.12)	(3.19)	(3.20)	(-1.42)