2018

Two Applications of Financial Economics to Real Estate

Jeffrey Dibartolomeo
University of Central Florida

TWO APPLICATIONS OF FINANCIAL ECONOMICS TO REAL ESTATE

by

JEFFREY A. DIBARTOLOMEO
B.S. Quinnipiac University, 2004
M.S. Florida Atlantic University, 2010

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Finance
in the College of Business Administration
at the University of Central Florida
Orlando, Florida

Spring Term

2018

Major Professor: Vladimir A. Gatchev

ABSTRACT

My first essay examines the effects of dividend policy on the liquidity risk of REITs. I argue that the mandatory high cash payouts of REITs reduce investor reliance on the stock market to satisfy their liquidity needs. Using a sample of equity REITs from 1980 through 2015, I find strong empirical evidence consistent with this paradigm. Unlike non-REIT property companies, I find REITs exhibit negative sensitivity to marketwide liquidity shocks; a result that is evident across most property type sectors. Moreover, while my findings are robust across a wide range of portfolios based on size, dividend frequency, leverage, market-to-book, operations type, and the presence of dividend reinvestment plans, smaller REITs mitigate liquidity risk only when their dividend frequency is relatively high. Finally, I find that price sensitivities to marketwide liquidity shocks increase after firms elect to discontinue REIT status. These findings strongly support the notions that investors view dividend payouts as a substitute for liquidity, and that REITs' relatively high mandated payout requirements benefit investors with reduced liquidity risk.

My second essay re-examines the ability of the Mills-Muth neoclassical land use theory to explain urban sprawl. I test the robustness of Brueckner and Fansler's (1983) seminal study using data drawn from the 1970 U.S. Census. A repeated sampling test shows that their 1970 sampling methodology led to spurious estimates; their conclusions regarding the economic factors driving sprawl cannot be supported. Nor can they be supported using more recent data from the 2000 and 2010 Census. Given this, I offer two alternate measures of urban sprawl: the traditional population density gradient and a new measure that relaxes the monotonicity constraint implied by traditional

density gradients. I find the factors identified by neoclassical theory better explain sprawl when using the density gradient and the non-monotonic measure than the Brueckner-Fansler approach.

For Helene and Anna.

ACKNOWLEDGMENTS

I would like to express my most sincere gratitude to my advisor, Dr. Vladimir Gatchev, and to the rest of my dissertation committee, Dr. Geoffrey Turnbull, Dr. David Harrison, Dr. Honghui Chen, and Dr. Charles Schnitzlein, for their invaluable mentorship, guidance, and support of my research. I am especially indebted to Dr. Gatchev, Dr. Turnbull, and Dr. Harrison for their patience, wisdom, and tutelage as I completed these two essays. I was very blessed to have you all serve as my committee, and I cannot thank you enough for leading me on this journey.

Additionally, I would like to thank non-committee members, Dr. Ajai Singh and Dr. Qinghai Wang, for their insightful suggestions and challenging questions in helping me refine this dissertation. I would also like to thank seminar participants at the 2015, 2016, and 2017 American Real Estate Society Annual Meetings for their thoughtful questions and comments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This dissertation comprises two essays on applications of financial economics to real estate. The first essay uses the unique dividend payout rules of real estate investment trusts (REITs) to examine the effects of dividend policy on the liquidity risk of REITs. REITs are required by law to distribute the vast majority of their ordinary taxable income as dividends. Given these regulatory mandates, as well as REITs' significant non-cash expenses such as depreciation, in practice it is not uncommon for REITs to have payout ratios in excess of 100%. Based on finance theory, I argue that the mandatory high cash payouts of REITs benefit investors by reducing investor reliance on the stock market to satisfy their liquidity needs. Miller and Modigliani (1961) predict a firm's payout policy is irrelevant and should not affect asset prices. However, their theory describes and requires a perfectly liquid capital market with no trading frictions. In practice, stock markets are not perfectly liquid and involve non-trivial frictions. To this point, Chordia, Roll, and Subrahmanyam (2000) find that liquidity has a common underlying component across different securities, leading to marketwide liquidity shocks, while Pástor and Stambaugh (2003) and Liu (2006) find significant evidence that this common component in liquidity is a state variable which is both material and value relevant with respect to asset prices.

Using a sample of equity REITs from 1980 through 2015, I find strong empirical evidence consistent with this paradigm along four key dimensions. First, unlike non-REIT property companies, REITs exhibit a negative sensitivity to marketwide liquidity shocks. More specifically, when marketwide liquidity declines, REIT prices tend to increase. Second, my findings are not property type specific, but rather are evident across broad classifications of property type sectors.

Interestingly, however, diversified REITs tend to provide less protection against adverse marketwide liquidity shocks than REITs with more focused investment strategies. Third, while my findings are robust across a wide range of portfolios based on REIT size, dividend frequency, leverage, market-to-book, operations type, and the presence of dividend reinvestment plans (DRIPs), smaller REITs provide protection to liquidity risk only when their dividend frequency is relatively high. Finally, examining only those firms electing to discontinue their REIT status, I find that price sensitivities to marketwide liquidity shocks increase after such changes. Taken together, the findings provide strong support for the notion that investors view dividend payouts as a substitute for liquidity, and further, that REITs, as a security class with relatively high regulatory mandated payout requirements, provide investors with an important benefit in the form of reduced liquidity risk.

The second essay re-examines the ability of the Mills-Muth neoclassical land use theory to explain urban sprawl. The policy debate over the extent to which urban sprawl is efficient or represents market failure has driven urban policy discussion since the start of the U.S. suburbanization process in the late 19th century (Mills and Hamilton, 1994). The neoclassical land use theory built on the seminal works by Mills (1967) and Muth (1969) provides a framework for understanding how real income growth, population growth, and long-run improvements in urban transportation technology drive the suburbanization of jobs and population typically identified as evidence of inefficient sprawl (Wheaton 1974, Bruckner 1987). Further, the debate over the extent to which urban sprawl represents market failure versus how much represents efficient market outcomes has been hampered by the lack of solid empirical analysis.

I begin empirical analysis in this essay by testing the robustness of Brueckner and Fansler's (1983) seminal study using data drawn from the 1970 U.S. Census. A repeated sampling test shows that their 1970 sampling methodology led to spurious estimates; their conclusions regarding the economic factors driving sprawl cannot be supported. Nor can they be supported using more recent data from the 2000 and 2010 Census. In light of this conclusion, this study offers two alternate measures of urban sprawl: the traditional population density gradient and a new measure, adapted from Wilder's (1978) financial market technical indicators, that relaxes the monotonicity constraint implied by traditional density gradients. While the factors identified by neoclassical theory cannot explain sprawl using the Brueckner-Fansler approach, this study finds that these factors do a better job explaining sprawl when using the density gradient and the non-monotonic measure than found previously in the literature.

ESSAY 1: THE EFFECT OF DIVIDEND POLICY ON LIQUIDITY RISK: EVIDENCE FROM REITS

1. Introduction

Real estate investment trusts (REITs) are required by law to distribute the vast majority of their ordinary taxable income as dividends. Given these regulatory mandates, as well as REITs' significant non-cash expenses such as depreciation, in practice it is not uncommon for REITs to have payout ratios in excess of 100%.[1] In comparison, non-REIT property companies are not required to pay any dividends on their common stock, and in recent years the aggregate payout ratio for stocks in the S&P 500 index has been substantively below that for REITs, with an average annual payout ratio hovering near 40%. From a valuation perspective, these relatively high REIT payout ratios offer the potential for significant clientele effects, while from a risk management perspective, the mandatory distribution requirements may increase and ensure liquidity for investors, reduce agency problems associated with managing and/or accounting for the firm's free cash flows, enhance the financial transparency of the organization, and as a result, reduce potential sources of investment risk for market participants.

To place these contentions more firmly within the context of the existing literature, consider first the implications strict dividend payout rules have for REIT prices. While the seminal work of Miller and Modigliani (1961) predicts a firm's payout policy is irrelevant and should not affect asset prices, their theory describes and requires a perfectly liquid capital market with no trading frictions. Within such a market, firms can costlessly repurchase shares to distribute free

[1] The REIT Modernization Act of 1999 reduced the mandatory minimum dividend distribution requirement to retain tax transparency for federal income tax purposes from 95% of ordinary taxable income to its current level of 90%.

cash flows, and investors with liquidity needs can costlessly sell shares and create homemade dividends. In practice, however, trading typically involves non-trivial frictions and stock markets are not perfectly liquid. This simple difference between theory and practice has potentially important ramifications, especially given the growing body of research, starting with Chordia, Roll, and Subrahmanyam (2000), finding that liquidity has a common underlying component. More specifically, the existing empirical evidence consistently shows liquidity is positively correlated across different securities, leading to marketwide liquidity shocks. Of great import along this dimension are the findings of Pástor and Stambaugh (2003) and Liu (2006), who provide significant evidence that this common component in liquidity is a state variable which is both material and value relevant with respect to asset prices.

Building upon these foundations, the current study contends that by paying dividends, firms provide an alternative source of liquidity to their investors. Within the confines of commercial property and REIT markets, I therefore argue that, due to their unique dividend payout requirements and subsequently high cash payouts, REITs benefit investors by reducing investor reliance on the stock market to satisfy their liquidity needs. This benefit should be especially evident and pronounced during marketwide liquidity shocks, as REITs have the potential to (at least partially) shield investors from such shocks. Thus, I propose REITs benefit investors by reducing their exposure to marketwide liquidity risk.

To test and evaluate this hypothesis, I examine the liquidity risk of REITs over the period 1980 through 2015. More specifically, following Pástor and Stambaugh (2003) I control for the market, size, and value factors of Fama and French (1993), the momentum factor of Carhart (1997), and use sample REITs' price sensitivity to marketwide liquidity shocks (i.e. REITs' liquidity betas) as a measure of liquidity risk. Previewing the core results, I find the prices of REIT stocks exhibit a significantly negative sensitivity to marketwide liquidity shocks. More directly, when market liquidity declines REIT prices increase. In contrast to REIT prices, the prices of property company stocks, which are not governed by the same mandatory distribution requirements as REITs, exhibit a statistically insignificant sensitivity to marketwide liquidity shocks. These findings are robust to both analyses of individual REITs and analyses performed at aggregate REIT levels using REIT indices.

Additionally, I find that the documented negative liquidity betas of equity REITs are not concentrated in REITs investing in any particular property type, but rather are evident across the majority of equity REITs. That said, I do find diversified REITs tend to provide less protection against adverse marketwide liquidity shocks than REITs with more focused investment holdings. Finally, I identify a subsample of firms that do not retain their REIT status throughout the entire sample period and compare their observed liquidity risk across intervals during which the REIT tax status election is, and is not, in place. Using this more selective sample, and consistent with my conceptual arguments, I find the price sensitivity to marketwide liquidity shocks increases after firms choose to discontinue their REIT tax status election.

I further examine the robustness of the findings by estimating the liquidity risk of REITs for a wide range of portfolios based on REIT size, dividend frequency, leverage, market-to-book, operations type, and the presence of dividend reinvestment plans (DRIPs). For the majority of these portfolios, I find REITs exhibit a significant negative sensitivity to liquidity risk. However, I do find that small REITs with infrequent dividend payments provide substantially less protection to liquidity risk than larger REITs and/or REITs with frequent dividend payments.

Taken together, these findings provide strong support for the notion that dividend policy has a significant effect on liquidity risk, and further suggest that due (in part) to their high dividend payouts, REITs exhibit relatively low liquidity risk. The findings also indicate, that from the perspective of investors, REITs become a more attractive investment during periods in which marketwide liquidity is constrained. As such, this study contributes to the understanding of how financial policy, and specifically dividend payout policy, affects asset prices, relative risk levels, and (potentially) expected returns of REITs in relation to other assets.

The remainder of the paper is organized as follows. Section 2 presents a review of the relevant literature and develops empirically testable hypotheses. Section 3 describes the data, variables, and methodology employed throughout the paper. Section 4 presents the main findings, while section 5 presents findings from a series of additional robustness tests. Finally, Section 6 summarizes the key findings, discusses their implications, and offers concluding remarks.

2. Literature Review and Hypotheses Development

Pástor and Stambaugh (2003) develop a liquidity risk factor, hereinafter referred to as PS, which measures a stock's sensitivity to unexpected changes in marketwide liquidity. Using a broad sample of non-REIT stocks, they find liquidity risk to be a priced factor which is not subsumed by the traditional (MKT, SML, and HML) risk factors of Fama and French (1993). Conceptually, the PS measure relies on return reversals on day t that follow signed dollar volume on day *t-1*. To demonstrate the intuition behind this metric, consider a day with heavy selling induced by a large shock to marketwide liquidity demand. Under such a scenario, stocks with a relatively high sensitivity to marketwide liquidity shocks will experience greater transitory price declines than stocks with lower sensitivities. Consequently, securities with higher sensitivities should experience stronger return reversals than those with lower sensitivities. The reason a liquidity-sensitive stock experiences a transitory decline in value on days with heavy selling is that investors seeking liquidity are relegated to trade with the market maker due to the absence of buyers, and the market maker's bid price may well reflect a steep discount. An investor that owns a liquidity-sensitive stock will thus experience a drop in wealth, as other shareholders of the same firm meet their liquidity demands by selling shares. *Ex ante*, the investor may not know whether the drop in wealth will be temporary, which will be the case if the investor continues to hold the stock, or permanent, which will be the case if the investor is affected by the marketwide liquidity shock and thus needs to sell. Liquidity risk thus stems (in part) from the uncertainty in investor future wealth due to aggregate liquidity shocks.

A priori, there are multiple important reasons to expect that liquidity risk may be lower for REITs than for similarly situated non-REIT firms. Of note, Clayton and McKinnon (2000) find market makers faced reduced risk of trading against informed traders during the REIT boom of 1993 due to an increase in liquidity traders that more than compensated for the increase in institutional traders during that time. Their results suggest that during hot markets, even on low-liquidity days, traders wishing to sell shares of a REIT may not find that the market maker is the only willing buyer. While the REIT boom has faded, and institutional investors now dominate the REIT market, both Blau, Nguyen, and Whitby (2015) and Ametefe, Devaney, and Marcato (2016) nevertheless report consistent and continuing improvements in REIT market liquidity over the past two decades.[2] Moreover, the general intuition in Pástor and Stambaugh (2003) can accommodate alternative sources of liquidity to investors, such as cash dividends. For example, Banerjee, Gatchev, and Spindt (2007) find dividends substitute for liquidity, and thus, investors do not necessarily have to create homemade dividends by selling shares. More specifically, using a sample of non-REIT stocks, they find stock sensitivities to liquidity shocks fall substantially after dividend initiations. However, for a typical common stock, dividend policy is a choice variable and is endogenously determined within the firm, thus making it difficult to establish a direct causal link between observed dividends and stock price sensitivities to liquidity shocks.[3] A focus on REITs provides a (partial) remedy to the endogeneity problem, as firms within this industry receive preferential tax treatment in exchange for meeting regulatory requirements including mandates regarding their sources of income and minimum dividend payout ratios. The requirement perhaps most pertinent to this paper is the regulatory mandate that REITs must disgorge a minimum of 90% of their ordinary taxable income as dividends in order to maintain tax transparency for Federal Income Tax purposes. Since investors in REITs have *ex ante* knowledge of this required minimum dividend payout ratio, they may well prioritize investing in REITs to enhance both their periodic cash flow returns and expected liquidity.

Conversely, many market participants and academic researchers argue the minimum dividend payout requirements represent a non-binding constraint for many, if not most, REITs. To

[2] See, for example, Cashman, Harrison, and Panasian (2016) for a discussion of the role of institutional investors in REIT markets.
[3] In contrast to Banerjee, Gatchev, and Spindt (2007), who completely exclude REITs from their estimation sample, the current investigation focuses extensively on these firms and exploits the legal environment of REITs to mitigate this endogeneity concern. Since REITs are required by law to distribute 90% of their taxable income as dividends, the decision to pay dividends is driven, in large part, by exogenous factors -- namely firm profits.

this point, nearly a quarter century ago Wang, Erickson, and Gau (1993) reported that many REITs pay out over 100% of their reported earnings. More recently, both Bradley, Capozza, and Seguin (1998) and Feng, Price, and Sirmans (2011) confirm this pattern continues to hold. While somewhat counterintuitive from traditional finance, accounting, or economics perspectives, REITs and other property companies can often sustain high payout ratios of this magnitude as their cash flows often greatly exceed net income due to large (non-cash) depreciation expenses.[4] Dividend payments in excess of regulatory minimums, up to 100% of taxable income, also offer potential tax advantages to the firm and its investors, as current period retained profits remain subject to federal income taxation at the corporate/entity level. For example, if a REIT distributes the 90% regulatory minimum, the firm will be taxed at corporate rates on the remaining 10%. Thus, REITs have an additional incentive to pay out more than the minimum to avoid an economic transfer of wealth from shareholders to the government, for which shareholders would likely punish the firm through a lower share price.[5] Conversely, large dividend payouts may come at the cost of leaving essentially no retained earnings to internally fund growth and expansion activities. However, Ott, Riddiough, and Yi (2005) suggest this cost may be relatively minor as REITs, on average, use internally generated capital to fund only 7% of their asset growth, while Hardin and Hill (2008) note the need for REITs to carefully manage their dividend policy decisions to ensure they convey meaningful information on the firm's operating position and expansion plans to key market

[4] Income and expense accruals, as well as capital raising activities may also support consistently high payout thresholds. Moreover, Hardin et al. (2009) show that the cash holdings of REITs are magnitudes lower than those of the average public firm.

[5] Other empirical evidence is also consistent with increased dividend payouts when taxation policy favors dividends. For example, Chetty and Saez (2005, 2006) find an increase in dividends after the Jobs and Growth Tax Relief Reconciliation Act of 2003 that reduced the maximum tax rate on dividends from 38% to 15%.

participants and other relevant stakeholders. Within this context, I derive specific, empirically hypotheses below.

Before outlining the main hypothesis, I make two important observations. First, the current investigation is not the only paper to investigate the interplay of REIT and non-REIT equities across equivalent trading venues. Notably, Subrahmanyam (2007) studies the joint dynamics between equity REIT and non-REIT stocks on the NYSE and finds effective spreads on non-REIT stocks forecast REIT spreads at both daily and monthly intervals. The author measures the participation of large (retail) traders by dollar (transaction) volume, and finds non-REIT order flow (measured in dollars) inversely leads REIT order flow (measured by transaction count). He subsequently attributes this relation to retail traders moving money into REITs when institutional traders sell non-REIT stocks. The current paper differs from Subrahmanyam (2007) both in its use of an aggregate measure of liquidity, and also by linking dividends to a stock's sensitivity to liquidity. Second, by construction, the PS liquidity factor explicitly excludes REITs. As the sample REITs trade on the same exchanges as the non-REIT property company stocks, the methodology employed in this paper measures the exposure of REITs to marketwide liquidity shocks in the markets for stocks that trade on these same venues.

Based upon the discussion above, the primary goal and contribution of the current manuscript is to examine and ascertain the relative importance of liquidity risk in REIT markets. More specifically, I advance the general proposition that, due to their unique dividend payout requirements, REITs reduce investor reliance on the liquidity provision role performed by secondary markets. The central hypothesis flows directly from this proposition, and may be summarized as follows:

Hypothesis 1: Investments in REITs are less sensitive to marketwide liquidity shocks than are investments in similarly situated property companies which have not elected REIT status.

Clearly, in addition to the focal regulatory considerations outlined above, the financial and operational transparency of a firm's activities may exert a significant influence over the firm's financial market liquidity. Across both the general finance and real estate investments literature, greater transparency has consistently been shown to enhance market based measures of liquidity, and lower investment risk along this key dimension.[6] In the work perhaps most relevant to the current investigation along this dimension, Danielsen and Harrison (2007) decompose the bid-ask spreads of REITs and find that the types of properties in which they invest materially impacts the liquidity of the firms' equity shares. Importantly, they find property type diversified investments lead to lower share liquidity. Conceptually, they argue diversification endows managers with real options for capital redeployment that investors find difficult to value. REITs with focused investment strategies do not confer such options to managers because of (self-imposed) constraints on the types of properties in which these firms invest. While the authors do not explicitly include non-REIT property companies in their analysis, I see no obvious reason why their results would not be generalizable to the broader cross section of real estate firms. Additionally, Capozza and Seguin (1999) find the market may penalize diversified REITs by way of reduced valuations, even though diversified REITs may not have lower cash flows. These findings again suggest diversified REITs may be viewed differently in the marketplace than their more property type focused peers. Given these previous findings, to examine whether the investment focus of both REITs and non-

[6] See, for example, Amihud, Mendelson, and Pedersen (2005), Lang and Maffett (2011), and the references therein.

REIT property companies materially influences the liquidity risk of these organizations, I examine and test *Hypothesis 2*:

Hypothesis 2: Property type diversified REITs exhibit a higher sensitivity to marketwide liquidity shocks than do their counterparts with property type focused investment strategies.

Even though all REITs are subject to the same payout rule, REITs differ in how frequently they provide dividends to their shareholders. While some REITs pay dividends every quarter, other REITs do not. There are two primary reasons for these observable differences. First, some REITs may simply elect to pay semi-annual dividends, which will naturally lead to a lower observed dividend frequency. Second, even if a REIT generally pays dividends every quarter, managers may decide to omit dividend payments in some quarters if, for instance, cash is not readily available. All else equal, investors relying on dividends to satisfy their liquidity needs will have a preference for more frequent dividend payments, or at the minimum avoid REITs that are likely to omit dividend payments. This consideration should be especially pronounced when the REIT is small, and hence when dividend omissions are likely to correspond to adverse economic conditions.[7] Based on these arguments, I propose and test the following hypothesis:

Hypothesis 3: Smaller REITs with infrequent dividend payments exhibit a higher sensitivity to marketwide liquidity shocks than larger REITs or REITs with frequent dividend payments.

Finally, as real estate markets are fraught with multiple sources of exogenous variation which may obscure key economic relations, and to provide a cleaner test of the focal relations, I

[7] The general finance literature consistently demonstrates firm size is highly correlated with the firm's level of financial constraint. See, for example, Kaplan and Zingales (2000).

lastly examine the subset of firms which experience a change in REIT status. Firms obtain REIT status by filing an election to do so with the U.S Internal Revenue Service and, in general, are free to alter this election at any time. As I seek to address whether a firm's REIT status is materially tied to their sensitivity to marketwide liquidity shocks (i.e., liquidity risk), unique insight may be gained from the examination of those firms which have operated under both REIT and non-REIT regulatory regimes. Consistent with the arguments advanced above, I propose the following hypothesis:

Hypothesis 4: A firm becomes more sensitive to liquidity shocks after it elects to discontinue its REIT status.

3. Data and Methodology

3.1 Data

In assembling the data necessary to conduct this investigation, I begin by identifying a sample of publicly traded equity REITs and non-REIT listed property companies over the period 1980 through 2015.[8] Regulatory changes regarding dividend payouts in the REIT industry over this period should not materially influence the test results, as while the REIT Modernization Act of 1999 reduced the dividend distribution requirement from 95% to 90% of ordinary taxable income, the key assertion that REITs and non-REIT property companies differ in their dividend payout policies due to dissimilar constraints and incentives remains valid both before and after the

[8] While I use this relatively long time frame for the core analyses, I also demonstrate that the findings are robust for both the earlier and more recent years of the sample.

passage and implementation of the act.[9] Since this paper relies on the *ex ante* knowledge that REITs must distribute a required percentage of their operating profits while no other real estate firm operates under such a restriction, non-REIT property companies provide an ideal matching sample of firms that operate within the real estate industry but do not face the exogenous dividend constraint of REITs.[10] As such, I sample all equity REITs in the CRSP/Ziman database and all firms identified as non-REIT property companies in the S&P Global Market Intelligence (formerly SNL Financial) database from January, 1980 through December, 2015. To ensure the findings are not driven by extremely small REITs, I exclude firms-years with a reported market capitalization of less than $20 million as measured in 2015 U.S. dollars.

For equity REITs and property companies in the sample, I next obtain stock market data over the sample period from CRSP, REIT characteristics from Ziman, and financial metrics from Compustat – Capital IQ. Unlike prior REIT studies that include only firms with positive income, I see no theoretical motivation to do so in this study, and only require firms to have non-missing dividend and earnings data. This identification procedure produces a sample of 440 unique equity REITs and 48 non-REIT property companies.[11]

---

[9] On the other hand, readers may reasonably wonder whether the rise of institutional investors throughout real estate markets which accompanied the onset of the modern REIT era in the early to mid-1990s alters the hypothesized relations. Reassuringly, unreported analyses conducted exclusively on both sample observations occurring post-1993, as well as on organizations incorporated subsequent to 1993, produce qualitatively similar results to those outlined below. These results should not be surprising, as while institutional investors have clearly changed the operating landscape of REIT markets, the dividend requirements which I aver are driving the mitigation of liquidity risk within this market sector are fundamentally independent of the presence of institutional investors in the marketplace. For additional insight on REIT market changes during the modern REIT era, see Ott, Riddiough, and Yi (2005).

[10] See 26 U.S.C §857 (2015) for the U.S. Internal Revenue code pertaining to REITs.

[11] On the surface, including firms with negative earnings that pay dividends leads to negative observed payout ratios. While obviously problematic from a conceptual perspective, dividend payout ratios do not enter directly into the estimation of the factor models, and thus do not pose serious econometric concerns. Rather, I contend the omission of firms with negative current period earnings would pose a greater threat to the generalizability of the focal results. As such, all available observations along this dimension are retained within the final estimation sample.

Table 1 reports financial and stock market characteristics for the sample firms, and highlights the comparability between REITs and non-REIT property companies. Among the important similarities are earnings and return on assets (ROA).[12] More specifically, the average ROA for the sample of equity REITs is 3.26%, while that of comparable non-REIT property companies averaged a nearly identical 3.24%. Given both sets of organizations operate within the same general business environment, and their performance is measured over an identical time horizon, this result is not surprising.[13]

Within the sample, equity REITs are significantly smaller than non-REIT property companies, when looking both at market capitalization and at asset size. Additionally, consistent with conventional wisdom, REITs are more likely to pay dividends, pay significantly more dividends, and have significantly higher payout ratios. Specifically, REITs pay dividends across 94% of the sample firm-years whereas property companies pay dividends in only 54% of the firm-years. Moreover, the average payout ratio of dividend-paying REITs is 2.36 times earnings. In contrast, the average payout ratio of dividend-paying non-REIT property companies is only 0.67 times earnings. Such differences are (at least partially) attributable to regulatory differences between the two firm types. Notably, the mandatory minimum distribution requirements prescribed for REITs virtually ensures their payout ratios will exceed those of their less regulated counterparts, while relatedly, the inability to retain profits and endogenously fund firm growth

[12] While REITs generally employ and report funds from operations (FFO) as their primary operating performance benchmark, to ensure compatibility with the matched non-REIT property companies the current investigation employs the more traditional return on assets (ROA) measure of accounting profitability.

[13] I readily recognize and cede the point that REITs, particularly those with strong growth plans or ambitions, maintain incentives to manage earnings from a dividend policy perspective. More specifically, accounting practices which reduce reported taxable income will lower mandatory dividend payments for REITs, thus facilitating increased capital retention activities, which, in turn, may lower capital acquisition costs and enhance expected firm growth.

may dampen REIT expansion activities and contribute to their relatively smaller firm size. Viewed in this light, the somewhat higher market-to-book ratios of property companies (1.56) relative to REITs (1.22) is not particularly surprising.

Interestingly, I also find REITs exhibit higher average monthly stock returns than the comparison set of property companies. This difference in returns could be explained, at least in part, by differences in size and market-to-book ratios across the two types of firms. For example, smaller firms and firms with lower market-to-book (sample REITs) are generally expected to generate higher returns than larger firms and firms with higher market-to-book ratios (sample non-REIT property companies). Additionally, these differences in return may be due to variations over time in the number of observations within each subsample. The subsequent estimation approach and empirical models examining the returns of both REITs and property companies account for these possibilities.

Continuing, debt utilization ratios across the two firm types reveal REITs employ less leverage in their capital structure than non-REIT property companies. The averages show REITs have a total debt ratio of 52%, while non-REIT property companies have a total debt ratio of 61%. While the observed leverage of sample REITs and property companies are somewhat higher than would be observed for the cross section of all publicly traded firms, these sample averages are similar to findings of recent real estate capital structure studies and lend support to the contention that the tangible nature of real estate asset holdings increases the debt capacity of firms within this market sector. [14]

Table 2 reports aggregate annual payout ratios for equity REITs and non-REIT property companies, calculated from annual earnings and dividends over the period 1980 through 2015. Consistent with the main hypothesis that REITs enhance investor liquidity, across all sample years, REITs paid out more dividends as a percentage of earnings than their non-REIT property company counterparts. Even during the financial crisis years of 2007 through 2009, when earnings were very low or negative for many sample firms, aggregate REIT dividend payout ratios remained relatively strong, though both 2008 and 2009 witnessed non-trivial declines in the total dollar amount of dividends distributed. As a point of comparison, aggregate non-REIT property company dividends in 2009 fell even more precipitously, to less than half of their 2007 level.

Table 3 further outlines the sample attributes by reporting the exchange trading venue for both sample REITs and non-REIT property companies. The majority of both types of firms trade on specialist exchanges, with a much smaller proportion trading through market makers. Across the two subsamples, 64% (18%) of REITs and 58% (15%) of property companies trade on the NYSE (NYSE American), while 19% of REITs and 27% of property companies trade on NASDAQ.[15] The similarities in trading venues between equity REITs and non-REIT property

---

[14] See, for example, Feng, Ghosh, and Sirmans (2007), Boudry, Kallberg, and Liu (2010), or Harrison, Panasian, and Seiler (2011).

[15] While Danielsen and Harrison (2000) find differences in liquidity between REITs that trade on organized exchanges compared to those that trade over-the-counter, their use of a microstructure liquidity measure differs greatly from my operationalization of liquidity.

companies provide confidence that the findings are not a relic of differences in the market making environment for the two types of firms.

3.2 Variables

To minimize model misspecification and ensure the results are attributable to liquidity risk rather than other potential sources of value relevant, systematic risk, I employ several risk factors commonly used in asset pricing studies to explain REIT and property company returns. The first three control metrics I include are the Fama and French (1993) market (MKT), small-minus-big (SMB, or "size"), and book-to-market (HML, or "value") factors. The market factor represents the value-weighted excess return on all stocks in the CRSP database over and above the risk-free rate. The size and value factors are constructed from 6 portfolios formed on size and book-to-market equity ratios, as follows. Stocks are halved at the median market capitalization into small versus large stocks, and stocks are similarly split into the lowest and highest 30% (growth and value stocks, respectively) and middle 40% (neutral stocks) of book-to-market equity ratios. SMB is the return to a portfolio that is long the three small portfolios (small-value, - neutral, and -growth) and short the three large portfolios (large-value, - neutral, and -growth). HML is similarly computed, and is the return to a strategy that is long value portfolios (small- and big-value) and short growth portfolios (small- and big-growth). Fama and French (1993) find these three factors explain over 90% of common stock returns.

The Carhart (1997) momentum factor (UMD) is also commonly used in asset pricing studies. Like the MKT, SMB, and HML factors, UMD is also the return to a zero-investment portfolio, and operationally is computed as the return to a portfolio which is long the past 1-year winners (highest return) and short the past 1-year losers (lowest return).[16] The fifth risk factor I employ is the Pástor and Stambaugh (2003) liquidity factor (LIQ). The resultant liquidity betas I estimate along this dimension are the primary focus of this study. More precisely, the liquidity factor is the return to a zero-investment portfolio which is long the least liquid stocks and short the most liquid stocks, where liquidity is measured by the sensitivity of prices to unexpected changes in marketwide liquidity. The estimated coefficient measures how resilient a stock is to marketwide liquidity shocks. Such an aggregated measure of liquidity is different from the microstructure metrics frequently employed across the REIT literature, which tend to focus on some variant of the firm's bid-ask spread. The fact that the liquidity factor is based on the non-REIT stock market makes it ideally suited to test if REITs provide investors protection from reductions in broader, marketwide liquidity.

Figure 1 charts the growth over the sample period of a $1 investment in January, 1980 in the liquidity factor, the market factor, the CRSP/Ziman value-weighted REIT index, and a value-weighted property company index. The liquidity and market factors are akin to portfolios comprised of the liquidity risk and market risk premia, respectively. When marketwide liquidity is low, returns to the liquidity portfolio should decline. Examining the chart, the REIT index tends to hold its value in periods of declining liquidity, while the market factor and property company index lose value. For example, consider the period from early 2000 through 2002. During this period, not only did the tech stock bubble burst, but the September 11th terrorist attacks on the U.S. occurred. Both events were marked by broad selling, which led to severe declines in

[16] Consistent with the existing literature on momentum factors, the annualized UMD factor omits previous month (t-1) returns to avoid complications associated with short-term return reversals.

marketwide liquidity. The long-short portfolio that is the liquidity factor saw large drops in value, as did the market and property company portfolios. Interestingly, the REIT portfolio actually gained value during this period, largely making up for returns that lagged the market from 1989 until the turn of the century.

3.3 Methodology

To test the central hypothesis that REITs are less sensitive to marketwide liquidity shocks than non-REIT property companies, I employ Fama and French (1993) three-factor and Carhart (1997) four-factor frameworks, augmented with the Pástor and Stambaugh (2003) liquidity factor. Specifically, I estimate the following model:

$$R_{it} - R_{ft} = \alpha + \beta_{LIQ} LIQ_t + \beta_{MKT} MKT_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{UMD} UMD_t + \mu_{it} \quad (1)$$

where for firm i in month t, R_{it} is the monthly return and R_{ft} is the rate on 1-month Treasury bills. The explanatory variables are the Fama and French (1993) market, size, and value factors and *UMD* is the Carhart (1997) 1-year momentum factor. Of key interest across these models are the liquidity betas (β_{LIQ}) associated with the Pastor and Stambaugh (2003) liquidity factor. Throughout the remainder of the paper I refer to Equation 1 as the 5-factor model.

4. Main Findings

4.1 Firm-Level Results for REITs and Property Companies

Consistent with prior studies that examine the returns of REITs using factor models, I expect the estimated coefficients to satisfy the following conditions: $\beta_{MKT} > 0, \beta_{SMB} > 0, \beta_{HML} >$

0, and $\beta_{UMD} < 0$ (see, for example, Chen, Downs, and Patterson, 2012). With respect to new innovations and insight offered by the current investigation, *Hypothesis 1* suggests the estimated REIT liquidity betas, β_{LIQ}, should be negative.

Table 4 reports estimates of the five-factor model separately for REITs and for non-REIT property companies, where I estimate the model using both firm-month observations and using value-weighted portfolios. As the results of these two alternative specifications are nearly identical, throughout the remainder of this section I focus the discussion primarily on the firm-level estimates. As predicted, β_{LIQ} is negative and significant at the 1% level for REITs, and statistically insignificant for property companies. Examining the firm-level estimates, I see that β_{LIQ} is equal to -0.094, suggesting REIT returns increase by approximately 0.094% for every 1% decrease observed in the liquidity of the non-REIT stock market. For non-REIT property companies, the point estimate for β_{LIQ} is equal to -0.008, which is also statistically indistinguishable from zero. More importantly, the difference in estimates is equal to -0.086 and is highly significant (at the 0.01 level), strongly suggesting REITs provide much greater protection against marketwide liquidity shocks than their non-REIT property company counterparts. The β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD} estimates also are all highly significant and exhibit the expected sign patterns outlined above. These results further indicate that, on average, REITs are less sensitive than non-REIT property companies to the market, size, and momentum factors. No differential sensitivity between REITs and property companies is observed with respect to the value factor.

The absolute magnitudes of the estimated betas reveal investor preferences for property companies along the market, size, and momentum factors and indifference between REITs and property companies when investor preference for value increases. This is not surprising given that property companies tend to be larger than REITs (as reported in Table 1). The property company factor model appears to exhibit somewhat better explanatory power than its REIT company counterpart, with an adjusted R-squared value of 20.92% for property companies compared to 19.98% for REITs.

4.2 Portfolio Results for REITs and Property Companies

Without further analysis, it is possible that the firm-level estimates are sample-dependent, or alternatively, that one of the restrictions imposed on the sample is driving the results. A simple test to help eliminate this possibility is to use portfolio, rather than firm, excess returns as the dependent variable. Consistent with the previous literature, I construct value-weighted portfolios using one-month lagged market capitalization as the value-weight of each REIT or property company.[17]

In general, the results closely follow the patterns previously reported for the firm-level estimates, and provide further evidence supporting the notion that REITs are less sensitive to non-REIT market liquidity shocks than a comparable sample of non-REIT property companies. The estimate for β_{LIQ} for REITs is again negative, and statistically significant at the 0.01 level. While still statistically insignificant, liquidity betas now exhibit an unexpected positive sign for property companies. More importantly, I again find that the difference is both negative (-0.127) and highly significant (at the 0.01 level).

[17] The use of lagged monthly market capitalizations to value-weight returns follows CRSP methodology for index construction.

Additionally, looking at the portfolio results, I now find no significant difference between the sensitivity of REITs and property companies to the size factor. However, I do find a difference in the value factor sensitivities, as REITs are now more sensitive to the value factor than property companies. This latter result is entirely consistent with the earlier findings that REITs tend to have somewhat lower market-to-book ratios than property companies. As a final comment, I also note that the portfolio models obtain substantially better fit, as signified by higher adjusted *R*-squared, than the firm-level models.

5. Robustness Analysis

5.1 Does Property Type Focus Affect Liquidity?

To examine the validity of *Hypothesis 2* and further explore the robustness of the results, I next examine whether the nature of the direct property investment affects equity REIT sensitivity to marketwide liquidity conditions. Of particular interest is whether the results of the previous section are concentrated within a few select types of property investments, or alternatively, whether the results are more general and apply to investments across the spectrum of property type holdings. Motivating this inquiry, Capozza and Seguin (1999) find that, among equity REITs, the shares of firms with more diversified property investments are less liquid than those of firms with more focused property investments. Additionally, they further note that while project-level cash flows tend to be higher for diversified REITs than for those with more property type focused investment portfolios, cash flows available to shareholders are roughly equivalent because of higher managerial and interest expenses. CRSP/Ziman reports property type index returns for diversified, health care, industrial/office, lodging/resorts, residential, retail, self-storage, and unclassified equity REITs. Using these property type value-weighted excess index returns over the risk-free rate as the dependent variables, I re-estimate the factor models and report the results in Table 6.[18]

Given the findings of Capozza and Seguin (1999), ex-ante I expect β_{LIQ} to be larger (i.e., less negative) for diversified equity REITs than for firms with investment holdings focused on a particular property type sector. Consistent with this expectation, I find β_{LIQ} is -0.076 (and significant at the 5% level) for diversified REITs. In contract, β_{LIQ} for REITs who specialize and focus their investments on health care, industrial and office, retail, and self-storage properties are all less than -0.100.[19]

Comfortingly, the return drivers for diversified REITs do not appear substantially different from those of more focused REITs with respect to any of the other traditional risk factors. Examining these additional factor loadings more closely, equity REITs invested in lodging and resort properties consistently exhibit the highest sensitivity to risk factors in the non-REIT stock market, with β_{MKT} of 0.968, β_{SMB} of 0.734, β_{HML} of 1.049, and β_{UMD} of -0.359, while equity REITs invested in healthcare and self-storage properties have among the smallest absolute factor loadings along each of these dimensions. Further to the previous point, the factor models for self-storage and health care REITs account for only 21.06% and 27.43% of the variation in their

[18] The number of monthly observation varies across property focus due to data availability from CRSP/Ziman. Nevertheless, index return series for most property types cover all months from 1980 through 2015.

[19] I note two exceptions to this general pattern. Specifically, the β_{LIQ} estimates for REITs that invest in either lodging/resorts and/or residential properties are higher than those found for diversified REITs. As such, I urge caution when drawing definitive conclusions regarding the relative magnitude of liquidity betas across these sectors.

respective property type indices, while the factor models for all other property type classifications exhibit significantly higher explanatory power with adjusted *R*-squared values ranging from approximately 42% to 51%.

5.2 Robustness over Time

I next examine and demonstrate the robustness of the estimates over time. The interest in variations of REIT liquidity risk over time is motivated by two considerations. First, the REIT Modernization Act of 1999 directly altered the dividend distribution requirements of REITs, and thus may have affected both the pricing of firms within this industry, as well as their sensitivity to liquidity risk. Given that my theoretical rationale applies both before and after this legislative change, I would like to ensure the findings are robust to the passage and implementation of the Act.[20] Second, the REIT Improvement Act of 2003 eliminated important tax disadvantages that served as effective barriers to REIT investment for non-U.S. citizens. These changes potentially increased the demand for publicly traded REIT equities across U.S. markets, thereby enhancing both market depth and REIT liquidity.[21]

To examine these issues, I re-estimate the model separately for two sub-periods: pre-1999 and post-2003. Panel A reports the estimates using firm-month level returns. Examining these estimates, consistent with the full sample analyses I find that the liquidity betas of REITs are significantly negative both in the pre-1999 sample and in the post-2003 periods. More specifically,

[20] In addition to altering mandatory dividend requirement, the Act dramatically broadened the array of services REITs could directly provide to their tenants. I make no a priori assumptions as to the anticipated effects of these additional changes on REIT liquidity betas.
[21] While the REIT Improvement Act of 2003 was initially introduced in Congress by James "Jim" McCrery III (R – Louisiana) on April 30th, 2003, it wasn't formally signed into law until October 2004 by President George W. Bush.

for the pre-1999 period, REITs exhibit a liquidity beta of -0.071. This implies that for every 1% decrease in marketwide liquidity, REIT prices increased by approximately 0.07%. While these subsample results suggest the focal relations are generally robust to the passage and implementation of the REIT Modernization Act, as the regulatory changed reduced mandatory dividends for REITs I are somewhat surprised to find liquidity betas are marginally higher after its implementation. To further examine the effects of regulatory changes on the REIT liquidity metrics, I next re-estimate the model exclusively on these firm-month observations following the introduction of the REIT Improvement Act of 2003. Interestingly, and entirely consistent with my central thesis that REIT investors care meaningfully about liquidity risk concerns, the estimated REIT liquidity betas and results regarding observable firm sensitivities to market wide liquidity shocks are even more pronounced during this latter, less restrictive era. Specifically, after 2003, the estimated REIT liquidity beta is -0.176, implying that for every 1% decrease in marketwide liquidity, REIT prices increase by approximately 0.176%. I therefore conclude that in more recent years, subsequent to the passage of key changes in REIT regulatory provisions, REIT prices have become more countercyclical with respect to marketwide liquidity risk. As an interesting aside, when I examine the estimates for non-REIT property companies, I again find property companies exhibit a significantly higher liquidity risk than REITs, and further, this difference is significant across both the pre-1999 and post-2003 periods. For the pre-1999 period, the difference in liquidity betas between the two types of firms is -0.139, while for the post-2003 period the difference is -0.099; both of these differences are statistically significant at the 0.01 level.

In Panel B, I re-estimate the model for these two sub-periods using the value-weighted portfolio return measured at a monthly frequency as the dependent variable. Not surprisingly, these portfolio estimates are similar to the firm-level estimates. More importantly, I again find that REITs have a lower liquidity risk than non-REIT property companies, and once again, these differences are statistically significant (at the 0.05 level for the pre-1999 sample, and at the 0.10 level for the post-2003 period).

5.3 Dividend Frequency and the Liquidity Risk of REITs

As mentioned before, REITs differ in how frequently they provide dividends to their shareholders. For example, of the 39,546 firm-months with available dividend data over the past eight quarters, I find that approximately 80% pay dividends in each of the past eight quarters. Thus, approximately 20% of sample firm-month observations come from organizations which paid dividends in fewer than eight quarters. Infrequent dividend payments may be of relatively little concern for larger, and hence less risky, REITs. On the other hand, as stated in *Hypothesis 3*, for small REITs, infrequent dividend payments may be a point of serious concern, particularly for liquidity-sensitive investors.

To test this idea, I re-estimate the five-factor model for double-sorts based on REIT size and dividend frequency. I thus construct six portfolios, where each month REITs are sorted independently into three portfolios based on firm size and two portfolios based on dividend frequency. I use the 30th and 70th percentiles of one-month lagged REIT market capitalizations to construct the three size portfolios. REITs are further classified into either paying dividends in each of the past eight quarters (dividend frequency equals 100%), or not paying dividends in all of the past eight quarters (dividend frequency less than 100%). In constructing these portfolio sorts, it is important to ensure that sample REITs exist for at least two years (eight quarters) to avoid potential misclassification issues. Thus, I estimate the model on only those REITs with available asset data over the past two years. The estimates for these portfolios are reported in Table 7.

Before discussing the estimates, I first note that dividend frequency is positively related to REIT size. For instance, within the less-than-100% dividend frequency portfolios, there are more small REITs than large REITs, whereas for the 100% dividend frequency portfolios, there are more large REITs than small REITs. These patterns further underscore the importance of controlling for REIT size when performing the analyses.

Turning to the estimates presented in Table 7, I find they are generally consistent with *Hypothesis 3*. Specifically, I find small REITs with a less-than-100% dividend frequency exhibit an estimated liquidity beta (β_{LIQ}) equal to -0.003, which is statistically indistinguishable from zero. This finding suggests small REITs that do not pay frequent dividends fail to provide protection against adverse marketwide liquidity shocks. This result is again entirely consistent with the idea that liquidity-sensitive investors discount the prices of such REITs when marketwide liquidity deteriorates. In contrast, small REITs with a 100% dividend frequency exhibit an estimated liquidity beta (β_{LIQ}) equal to -0.107, which is statistically significant at the 0.01 level. Moreover, the two estimates are significantly different from each other at the 0.01 level.

When I examine the liquidity sensitivities of medium-sized REITs, I find that β_{LIQ} is similar across REITs regardless of their dividend frequency (-0.102 versus -0.117). For large REITs, I find those firms in the 100% dividend frequency portfolio have a less negative liquidity beta (β_{LIQ} = -0.095) than do firms in the less-than-100% dividend frequency portfolio (β_{LIQ} = -0.173). To further investigate this intriguing result, I next examine REIT size for the two large

REIT portfolios. Interestingly, I find that the average REIT size is substantially higher in the less-than-100% dividend frequency portfolio than in the 100% dividend frequency portfolio, suggesting the portfolio sorts may not perfectly control for REIT size. In an attempt to address this problem, I bifurcate the less-than-100% dividend frequency portfolio of large REITs into two additional portfolios: large-1 and large-2. After doing so, I find that the highly negative estimate of β_{LIQ} for the less-than-100% dividend frequency portfolio comes from the largest REITs in the sample (i.e., the ones entering the large-2 portfolio). In fact, the estimate of β_{LIQ} for the large-1 portfolio of REITs with less-than-100% dividend frequency is indistinguishable from the estimates for β_{LIQ} for the large portfolio of REITs with 100% dividend frequency.

Taken together, the evidence presented in this section is consistent with the idea that investors value REITs, at least in part, based on their ability to deliver liquidity through dividends. Specifically, I find REIT prices respond less favorably to adverse marketwide liquidity shocks when REITs are less likely to meet investor liquidity needs, such as in the case of smaller REITs with a lower frequency of past dividend payments.

5.4 The Liquidity Risk of REITs Conditional on Firm Characteristics

In further robustness tests I examine the liquidity risk of REIT portfolios formed based on various double-sorts using firm size, leverage, market-to-book ratios, operations type, and the presence or absence of a dividend reinvestment plan. My main objective in performing these analyses is to ensure the main findings are not concentrated within a subset of REITs with particular financial characteristics. With that objective in mind, I report and focus my discussion on the estimates of the resultant liquidity betas.[22]

Table 8 presents the estimates of β_{LIQ} from Equation (1) for double-sorts based on both firm leverage and market-to-book ratios. I thus estimate the model for nine portfolios, where each month REITs are sorted independently into three portfolios based on leverage and three portfolios based on market-to-book. To construct these portfolios, I use the 30th and the 70th percentiles of lagged REIT leverage (total liabilities over total assets) and lagged market-to-book of assets (the sum of market capitalization and total liabilities divided by total assets).

Examining the estimates in Table 8, I do not find any particular patterns that consistently relate β_{LIQ} to firm leverage or market-to-book. For instance, the least negative β_{LIQ} (-0.050) is estimated for the portfolio of firms with low leverage and low market-to-book, while the most negative β_{LIQ} (-0.125) is estimated for the portfolio of firms with medium leverage and again low market-to-book (both estimates are statistically significant at the 0.01 level, but are not significantly different from each other). The finding that liquidity risk is similar across REITs with different leverage suggests that, on average, REITs set their capital structure at levels that do not adversely affect their ability to generate and distribute earnings to shareholders.

Table 9 presents the estimates of β_{LIQ} for double-sorts based on REIT size and operations type. Specifically, I estimate the model for six portfolios, where each month REITs are sorted independently into three portfolios based on size and two portfolios based on REIT operations type. I use the 30th and 70th percentiles of one-month lagged REIT market capitalizations to

[22] All estimates are available from the authors upon request.

construct the three size portfolios, while REIT operations type is further classified into either development or operational based upon whether or not the firm has recently reported the existence of any new projects in their development pipeline.

Interestingly, I find that development REITs have a significantly lower liquidity risk than operational REITs. For example, β_{LIQ} of small/operational REITs is equal to -0.076 whereas β_{LIQ} of small/development REITs is equal to -0.179. The difference between the two estimates is both economically and statistically significant (at the 0.01 level). I find similar difference patterns for both medium-sized and large REITs as well. One possible explanation for these findings is that REITs with relatively low (high) risk levels – including exposure to liquidity risk – are rewarded (penalized) with a lower (higher) cost of capital, and are thus more (less) likely to undertake new development projects. Regardless of the root cause of these differences, consistent with my focal hypotheses, I note that even operational REITs exhibit a significantly negative sensitivity to marketwide liquidity shocks.

Finally, Table 10presents the estimates of β_{LIQ} for double-sorts on both firm size and the presence of a dividend reinvestment plan. I estimate the model for six portfolios, where each month REITs are sorted independently into three portfolios based on firm size and two portfolios based on whether or not the REIT offers a dividend reinvestment plan. I again use the 30th and 70th percentiles of one-month lagged REIT market capitalizations to construct the three size portfolios. Dividend reinvestment plans allow investors to reinvest dividends back into the issuing company without incurring significant transactions costs. However, such a benefit will only be of unique relevance: 1) to investors that do not consider dividends as supplying their liquidity needs, and 2) only if those same investors face relatively high transaction costs in reinvesting such dividends through a broker.

Examining the results, I do not find any discernable differences in liquidity betas between REITs that have a dividend reinvestment program and those that do not. For instance, small REITs with a dividend reinvestment program have a β_{LIQ} of -0.082, which is quite similar to (and statistically indistinguishable from) the -0.100 estimate of β_{LIQ} for small REITs with no dividend reinvestment program in place. As such, I conclude that while REITs with dividend reinvestment programs evidence a slightly lower liquidity risk than comparable firms without such a program, the differences are relatively trivial. While these estimates could be viewed as evidence that dividend reinvestment programs further enhance the liquidity of REITs (especially for investors without pressing liquidity needs or concerns), the small differences in estimates suggest the lack of a dividend reinvestment program does not impose material illiquidity costs on investors.

Overall, the findings presented in this section demonstrate that REITs provide significant protection against marketwide liquidity shocks regardless of REIT size, leverage, market-to-book, REIT operations type, and/or the presence of dividend reinvestment plans. The similar findings across different portfolios are once again consistent with my focal hypothesis that it is the legal status of REITs (and not their particular financial characteristics) that enhances investor liquidity.

5.5 Does REIT Status Affect Liquidity?

Having established that equity REITs offer market participants an investment with a relatively low sensitivity to marketwide liquidity shocks, I next examine what happens to the shares of a firm that elects to discontinue its REIT status. If the previously described liquidity risk results are the by-product of REIT regulations which credibly commit firms to high dividend payments, when a firm discontinues this special tax status election and becomes a taxable property company, investors may no longer be willing or able to rely on the continuation of such robust dividend payouts. As a consequence, the shares of the company may become more sensitive to marketwide liquidity shocks. Alternatively, to the extent the previous results are driven by firm characteristics and attributes not sufficiently controlled for in previous model specifications, the change in tax status election should contain little to no information regarding the firm's future dividend payouts, and thus, such changes should not influence the firm's sensitivity to marketwide liquidity shocks.

To explore these conflicting hypotheticals and to provide a test of *Hypothesis 4*, I augment the factor models of Equation (1) with an indicator variable Status, which takes the value of one if the firm is operating as a REIT at the time of the observation, and zero if the firm has not elected REIT status for that period.[23] The estimation sample for this final robustness test includes only those firms that operated as a REIT at some point during the sample observation period. CRSP/Ziman reports both the date a firm first obtains REIT status and when the firm discontinues this election. Using this information, I identify those months during which each firm operates as a REIT and as a non-REIT property company. The augmented factor model used to investigate the importance of this selection is specified as follows:

$$R_{it} - R_{ft} = \alpha_1 Status_{it} + \alpha_0(1 - Status_{it}) + \beta_1 Status_{it} \times F_t + \beta_0(1 - Status_{it}) \times F_t + u_{it} \quad (2)$$

[23] One potential concern with the results of this subsection is that many REITs who discontinue their REIT status do so in response to a merger, acquisition, or bankruptcy. Frequently, this leaves only one monthly observation with return data for out-of-status REITs, thereby limiting the power of the empirical tests of this relation.

In this model, F_t is a 5×1 vector containing all five previously employed risk factors, while the β vectors contain the corresponding coefficients for the five factors. This specification allows estimation of the factor sensitivities of returns both when companies have elected REIT status (β_1) and when they have not (β_0).

The estimates from this augmented factor model are reported in Table 11. The results are once again consistent with the notion that REIT firms are relatively insulated from marketwide liquidity shocks. The liquidity beta (β_{LIQ}) for stocks operating under the REIT status election equals -0.096. In contrast, when these same firms do not operate as REITs their liquidity beta (β_{LIQ}) is equal to -0.025, and is statistically indistinguishable from 0. More importantly, the difference in liquidity betas across this choice of firm status has a *t*-statistic of 2.16, and is thus statistically significant at the 0.05 level. These findings highlight the differential liquidity risk faced by REITs versus non-REIT property companies. More specifically, when property companies elect to not retain their existing REIT status, their shares become more sensitive to marketwide liquidity shocks.

In sum, the results reported in this subsection support the conclusion that liquidity is directly tied to a firm's status as a REIT, and moreover, reinforce the prior conclusion that equity REITs have lower sensitivities to marketwide liquidity shocks than similarly situated non-REIT property companies.

6. Concluding Remarks

The primary purpose of this paper is to determine the sensitivity of REIT returns to unexpected changes in marketwide liquidity. The paper further investigates the types of REITs, by investment focus, that are most sensitive to marketwide liquidity shocks. In completing this study, I use a comparison sample of both REITs and non-REIT property companies that do not face the same regulatory mandate to pay high dividends. The main findings may be summarized as follows. First, equity REITs, as a group, exhibit a negative sensitivity to marketwide liquidity shocks, such that when marketwide liquidity declines, REIT values increase. I interpret this finding as providing evidence that REIT prices reflect a liquidity benefit to investors. In contrast, non-REIT property companies exhibit no such relation, as their prices do not increase when marketwide liquidity deteriorates. Together, these findings suggest investors view the high mandatory dividend payouts of REITs as a substitute for liquidity, especially when marketwide liquidity is low.

Second, the documented effects are not specific to REITs investing in any single property type category, but rather appear to influence the pricing of equity REITs across the spectrum of alternative property type investment holdings. Prior studies have found that both diversified REITs, and those invested in difficult to value property types, suffer from lower valuations and reduced liquidity. I similarly find that diversified REITs, as well as those invested in lodging, resort, and residential properties, exhibit higher sensitivities to marketwide liquidity shocks than firms with more property type focused investment holdings, and/or whose investments are concentrated in more easily valued assets. Third, the liquidity benefits of REITs do not extend to small REITs when such firms have infrequent dividend payments. I attribute this result, in part, to the inability of such firms to credibly commit to sustaining dividend levels which liquidity conscious investors value. Lastly, as a fourth and final test of my focal hypothesis, I examine how the liquidity betas of stocks change when firms elect to drop their REIT status and become fully taxable, non-REIT property companies. Consistent with the primary findings, I again observe firm share price sensitivities to marketwide liquidity shocks increase when companies elect to discontinue their REIT status.

In conclusion, the findings presented in this paper are consistent with the notion that the structure of financial claims may exert a significant impact on the values of those claims. In particular, due to market imperfections, the payout structure of equity securities, such as REITs, is highly relevant. Finally, the findings further suggest that REITs, as a security class with demonstrably high regulatory mandated payout requirements, provide investors with an important benefit in the form of reduced exposure to liquidity risk.

References

Ametefe, F., S. Devaney, and G. Marcato. 2016. Liquidity: A Review of Dimensions, Causes, Measures, and Empirical Applications in Real Estate Markets. Journal of Real Estate Literature 24(1): 1–29.

Amihud, Y., H. Mendelson, and L. Pedersen. 2005. Liquidity and Asset Prices. Foundations and Trends in Finance 1, 269–364.

Banerjee, S., V.A. Gatchev and P.A. Spindt. 2007. Stock Market Liquidity and Firm Dividend Policy. Journal of Financial and Quantitative Analysis 42: 369–397.

Blau, B.M., N. Nguyen, and R.J. Whitby. 2015. The Distribution of REIT Liquidity. Journal of Real Estate Literature 23(2): 233–252.

Boudry, W.I., J.G. Kallberg, and C.H. Lius. 2010. An Analysis of REIT Security Issuance Decisions. Real Estate Economics 38(1): 91–120.

Bradley, M., D.R. Capozza, and P.J. Seguin. 1998. Dividend Policy and Cash-Flow Uncertainty. Real Estate Economics 26(4): 555–580.

Capozza, D.R. and P.J. Seguin. 1999. Focus, Transparency and Value: the REIT Evidence. Real Estate Economics 27: 587–619.

Carhart, M.M. 1997. On Persistence in Mutual Fund Performance. The Journal of Finance 52: 57–82.

Cashman, G.D., D.M. Harrison, and C.A. Panasian. 2016. Clawback Provisions in Real Estate Investment Trusts. Journal of Financial Research 39(1): 87-114.

Chen, H., D.H. Downs and G.A. Patterson. 2012. The Information Content of REIT Short Interest: Investment Focus and Heterogeneous Beliefs. Real Estate Economics 40: 249–283.

Chordia, T., R. Roll and A. Subrahmanyam. 2000. Commonality in Liquidity. Journal of Financial Economics 56, 3–28.

Clayton, J. and G. MacKinnon. 2000. Measuring and Explaining Changes in REIT Liquidity: Moving Beyond the Bid–Ask Spread. Real Estate Economics 28, 89–115.

Danielsen, B. and D. Harrison. 2000. The Impact of Potential Private Information on REIT Liquidity. Journal of Real Estate Research 19: 49–71.

Danielsen, B. and D. Harrison. 2007. The Impact of Property Type Diversification on REIT Liquidity. Journal of Real Estate Portfolio Management 13: 329–344.

Fama, E.F. and K.R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics 33: 3–56.

Feng, Z., C. Ghosh, and C.F. Sirmans. 2007. On the Capital Structure of Real Estate Investment Trusts (REITs). Journal of Real Estate Finance and Economics 34(1): 81–105.

Feng, Z., S.M. Price, and C.F. Sirmans. 2011. An Overview of Equity Real Estate Investment Trusts (REITs): 1993-2009. Journal of Real Estate Literature 19(2): 307–343.

Hardin III, W. and M.D. Hill. 2008. REIT Dividend Determinants: Excess Dividends and Capital Markets. Real Estate Economics 36: 349–369.

Hardin III, W., M.J. Highfield, M.D. Hill, and W. Kelley. 2009. The Determinants of REIT Cash Holdings. Journal of Real Estate Finance and Economics 39(1): 39–57.

Harrison, D.M., C.A. Panasian, and M.J. Seiler. 2011. Further Evidence on the Capital Structure of REITs. Real Estate Economics 39(1): 133–166.

Kaplan, S.N. and L. Zingales. 2000. Investment Cash-flow Sensitivities are Not Valid Measures of Financing Constraints. Quarterly Journal of Economics 115(2): 707-712.

Lang, M., and M. Maffett. 2011. Transparency and Liquidity Uncertainty in Crisis Periods. Journal of Accounting and Economics 52: 101–125.

Liu, W. 2006. A Liquidity-augmented Capital Asset Pricing Model, Journal of Financial Economics 82: 631–671.

Miller, M.H., and F. Modigliani. 1961. Dividend Policy, Growth, and the Valuation of Shares. Journal of Business 34: 411–433.

Ott, S.H., T.J. Riddiough and H.C. Yi. 2005. Finance, Investment and Investment Performance: Evidence from the REIT Sector. Real Estate Economics 33: 203–235.

Pástor, L. and R.F. Stambaugh. 2003. Liquidity Risk and Price Discovery. Journal of Political Economy 111: 642–685.

Subrahmanyam, A. 2007. Liquidity, Return and Order-Flow Linkages Between REITs and the Stock Market. Real Estate Economics 35: 383–408.

Wang, K., J. Erickson and G.W. Gau. 1993. Dividend Policies and Dividend Announcement Effects for Real Estate Investment Trusts. Real Estate Economics 21: 185–201.

Figures and Tables

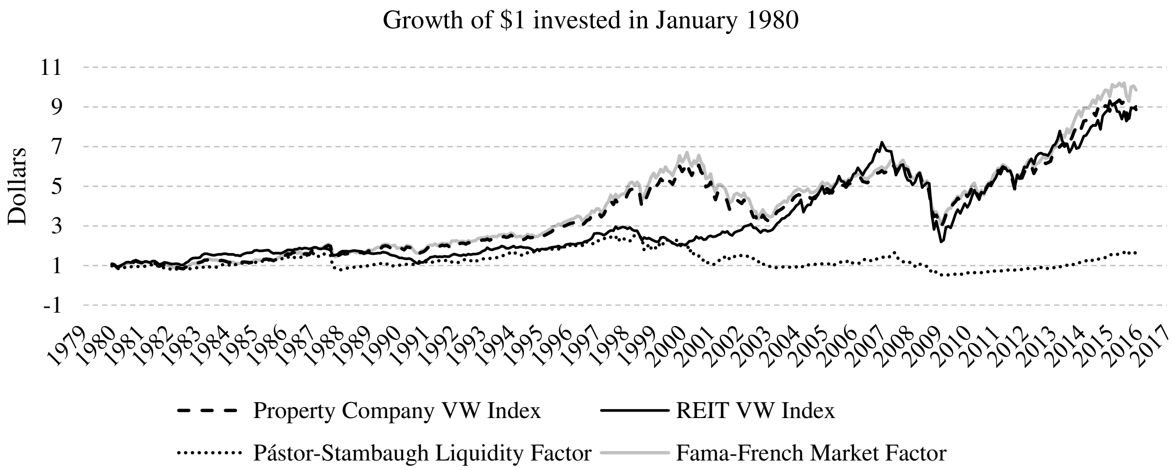

Figure 1: Growth of Portfolios

This figure plots the growth over time of a one-dollar investment in the Pástor and Stambaugh (2003) liquidity factor, the Fama and French (1993) market factor, the CRSP/Ziman value-weighted REIT index, and a value-weighted property company index. The liquidity factor is the return to a portfolio which is long the least liquid stocks and short the most liquid stocks. The market factor is the excess value-weighted return on stocks over the 1-month Treasury bill rate. The property company index is the excess return over and above the 1-month Treasury bill rate to a value-weighted portfolio of all property companies in the sample.

Table 1: Summary Statistics

This table reports summary statistics for equity REITs (Panel A) and property companies (Panel B), as well as Wilcoxon rank-sum tests of differences between the two sets of firms (Panel C). I identify 440 equity REITs from the CRSP/Ziman database and 48 property companies from the S&P Global Market Intelligence (formerly SNL) database with return data in the Center for Research in Securities Prices (CRSP) monthly files over the interval 1980 through 2015. I exclude firm-years with a market capitalization below $20 million. Compustat provides data on annual assets, non-diluted earnings per share (EPS) excluding extraordinary items, dividends, net income, the number of shares outstanding, the number of shares used to calculate non-diluted EPS, total liabilities, and quarterly dividends. REIT month-end market capitalization is obtained from the CRSP/Ziman database. Property company market capitalization is the month-end closing stock price multiplied by the number of shares outstanding. Earnings are calculated as the non-diluted EPS excluding extraordinary items multiplied by the number of shares used to calculate non-diluted EPS. Market capitalization, assets, earnings, and dividends are reported in millions of 2015 U.S. dollars using the GDP deflator from the U.S. Bureau of Labor Statistics. The dividend payer dummy is set equal to one if a firm reports a positive dividend in a given year, and zero otherwise. The dividend payout ratio is calculated as dividends divided by earnings for firm-years with positive earnings and positive dividends. Return on assets (ROA) equals net income divided by total assets, leverage equals total liabilities divided by total assets, and the market-to-book of assets is the ratio of a firm's market value to book value of total assets, where firm market value is the sum of the firm's market capitalization and total liabilities. Dividend frequency is the moving average of the past eight quarters of a binary variable that equals one if the firm paid positive dividends in a quarter, and zero otherwise.

Panel A: REITs (4,014 firm-year observations)

	Mean	Median	Standard deviation	25^{th} percentile	75^{th} percentile
Market capitalization (mill. 2015 USD)	1,660.46	525.66	3,380.41	154.97	1,676.95
Assets (mill. 2015 USD)	2,350.66	1,012.80	3,704.13	311.14	2,788.28
Earnings (mill. 2015 USD)	41.43	16.92	112.50	2.57	51.69
Dividends (mill. 2015 USD)	77.57	34.39	121.64	10.05	91.47
Dividend payer dummy	0.94				
Payout ratio for dividend payers	2.36	1.32	4.72	1.02	2.04
ROA (%)	3.26	2.92	8.64	1.12	4.92
Monthly return (%)	1.15	1.02	7.92	– 2.70	4.97
Leverage	0.52	0.53	0.21	0.41	0.64
Market-to-book of assets	1.22	1.14	0.59	0.96	1.36
Dividend frequency (8-qtr. moving avg.)	0.91	1.00	0.23	1.00	1.00

Table 1: Summary Statistics

Panel B: Property companies (579 firm-year observations)

	Mean	Median	Standard deviation	25th percentile	75th percentile
Market capitalization (mill. 2015 USD)	1,935.25	361.34	3,579.73	115.02	2,128.71
Assets (mill. 2015 USD)	2,454.30	721.18	4,292.65	329.72	2,180.26
Earnings (mill. 2015 USD)	65.06	13.64	196.27	– 2.26	74.97
Dividends (mill. 2015 USD)	23.53	0.89	60.59	0.00	17.72
Dividend payer dummy	0.54				
Payout ratio for dividend payers	0.67	0.34	1.86	0.17	0.55
ROA (%)	3.24	2.96	7.28	– 0.12	5.94
Monthly return (%)	0.99	0.51	10.70	– 4.88	6.42
Leverage	0.61	0.61	0.24	0.47	0.77
Market-to-book of assets	1.56	1.19	1.21	0.95	1.73
Dividend frequency (8-qtr. moving avg.)	0.44	0.38	0.43	0.00	1.00

Panel C: Wilcoxon rank-sum tests of differences

	Difference in means (REITs – property companies)			Difference in medians (REITs – property companies)		
	Diff.	z-score	p-value	Diff.	z-score	p-value
Market capitalization (mill. 2015 USD)	– 274.79	5.57	0.001	164.32	9.29	0.001
Assets (mill. 2015 USD)	– 103.64	1.91	0.056	291.61	4.66	0.001
Earnings (mill. 2015 USD)	– 23.63	0.90	0.366	3.28	1.55	0.121
Dividends (mill. 2015 USD)	54.04	21.85	0.001	33.50	15.42	0.001
Dividend payer dummy	0.40	29.35	0.001	0.00	29.35	0.001
Payout ratio for dividend payers	1.70	21.80	0.001	0.99	14.08	0.001
ROA (%)	0.00	0.71	0.475	0.00	0.05	0.960
Monthly return (%)	0.16	4.21	0.001	0.51	4.52	0.001
Leverage	– 0.10	9.24	0.001	– 0.08	7.43	0.001
Market-to-book of assets	– 0.34	4.70	0.001	– 0.05	1.83	0.068
Dividend frequency (8-qtr. moving avg.)	0.48	55.66	0.001	0.63	49.49	0.001

Table 2: Aggregate Earnings, Dividends, and Payout Ratios

This table reports aggregate earnings, dividends, and payout ratios for the REIT and property company samples during each year of the sample period: 1980 - 2015. Specifically, I identify 440 equity REITs from the CRSP/Ziman database and 48 property companies from the S&P Global Market Intelligence (formerly SNL) database with return data in the Center for Research in Securities Prices (CRSP) monthly files. I exclude firm-years with a market capitalization below $20 million. Compustat provides data on annual non-diluted earnings per share (EPS) excluding extraordinary items, dividends, and the number of shares used to calculate non-diluted EPS. Earnings are calculated as the non-diluted EPS excluding extraordinary items multiplied by the number of shares used to calculate non-diluted EPS. Earnings and dividends are reported in millions of 2015 U.S. dollars using the GDP deflator from the U.S. Bureau of Labor Statistics to adjust nominal values for inflation. The aggregate payout ratio is equal to aggregate dividends divided by aggregate earnings.

Table 2: Aggregate Earnings, Dividends, and Payout Ratios

Year	REITs				Property companies			
	Firms	Aggregate earnings	Aggregate dividends	Aggregate payout	Firms	Aggregate earnings	Aggregate dividends	Aggregate payout
1980	33	449.45	454.57	1.01	5	337.34	105.67	0.31
1981	38	311.23	272.56	0.88	4	312.01	113.00	0.36
1982	42	370.29	340.90	0.92	4	225.91	117.94	0.52
1983	40	548.94	480.04	0.87	4	253.63	129.64	0.51
1984	39	705.22	578.07	0.82	6	308.99	111.27	0.36
1985	43	983.92	1,111.58	1.13	4	286.56	281.69	0.98
1986	47	634.26	704.89	1.11	5	317.03	121.99	0.38
1987	57	757.85	947.77	1.25	7	473.34	200.31	0.42
1988	51	802.63	970.78	1.21	8	433.02	220.21	0.51
1989	44	741.55	876.93	1.18	7	403.25	227.32	0.56
1990	43	745.44	916.94	1.23	6	392.65	213.34	0.54
1991	57	684.07	936.65	1.37	8	309.95	219.58	0.71
1992	57	759.20	938.71	1.24	7	270.29	165.34	0.61
1993	90	1,283.33	1,667.75	1.30	8	235.91	159.67	0.68
1994	144	2,495.16	3,723.78	1.49	7	228.66	139.42	0.61
1995	149	3,319.79	4,635.43	1.40	6	247.87	146.12	0.59
1996	146	4,610.37	5,676.34	1.23	7	457.67	154.88	0.34
1997	152	6,399.25	8,307.69	1.30	9	500.40	205.62	0.41
1998	155	9,018.57	12,098.14	1.34	7	867.00	230.42	0.27
1999	154	10,666.40	12,492.49	1.17	8	1,070.03	147.23	0.14
2000	137	11,881.46	12,370.67	1.04	7	1,253.61	162.38	0.13
2001	115	8,658.22	11,052.60	1.28	6	771.45	154.31	0.20
2002	107	8,379.29	11,663.38	1.39	6	1,112.56	159.73	0.14
2003	104	8,608.20	12,355.40	1.44	6	1,036.51	272.42	0.26
2004	109	9,125.57	13,933.45	1.53	7	1,172.32	230.86	0.20
2005	105	8,713.64	13,627.69	1.56	10	1,540.98	287.48	0.19
2006	99	8,919.89	13,069.35	1.47	10	2,988.87	812.18	0.27
2007	86	9,539.59	12,613.02	1.32	11	2,396.38	516.86	0.22
2008	76	4,953.17	10,369.96	2.09	9	946.00	427.71	0.45
2009	62	4,202.72	7,109.47	1.69	7	763.87	198.47	0.26
2010	65	5,358.62	8,958.66	1.67	8	1,804.36	417.41	0.23
2011	70	6,481.41	10,062.21	1.55	9	2,149.51	492.56	0.23
2012	84	7,899.02	13,516.86	1.71	10	2,367.59	1,365.77	0.58
2013	101	9,696.95	16,403.07	1.69	11	3,020.91	952.80	0.32
2014	121	18,316.48	23,965.34	1.31	11	2,585.22	1,296.62	0.50
2015	129	21,237.01	26,830.85	1.26	12	4,079.91	1,040.83	0.26
Overall	440	198,258.16	276,034.00	1.39	48	37,921.55	12,199.03	0.32

Table 3: Listing Exchange

This table reports the listing exchange for each firm in the REIT and property company samples. I identify 440 equity REITs from the CRSP/Ziman database and 48 property companies from the S&P Global Market Intelligence (formerly SNL) database. I exclude firm-years with a market capitalization below $20 million. The listing exchange for each firm is obtained from the Center for Research in Securities Prices (CRSP). For firms that changed listing exchanges during the sample period, the table reports only the last exchange. The three listing exchanges represented in the sample are the New York Stock Exchange (NYSE) American (formerly NYSE MKT and the American Stock Exchange), the New York Stock Exchange (NYSE), and the National Association of Securities Dealers Automated Quotation system (NASDAQ).

	REITs		Property companies	
Exchange	Firms	% of Sample	Firms	% of Sample
NYSE American	77	18	7	15
NYSE	280	64	28	58
NASDAQ	83	19	13	27
Total	440	100	48	100

Table 4: The Liquidity Risk of REITs and Property Companies

This table reports OLS regression estimates from a five-factor model explaining the monthly excess returns of equity REITs and of property companies. The table also reports the differences in estimates between the two sets of firms. In the firm-level regressions, the dependent variable is each firm's monthly stock return in excess of the 1-month Treasury rate. I also estimate the model using as dependent variables the monthly value-weighted excess return of REITs and of property companies, where weights are based on each firm's past-month market capitalization. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). The coefficients for these factors are β_{LIQ}, β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD} respectively. The reported *t*-statistics (in parentheses) test whether estimates are significantly different from 0. [a], [b] and [c] denote significance at the 0.01, 0.05, and 0.10 levels.

The dependent variable is the monthly excess return

	Firm-level		Value-weighted portfolio		Differences in coefficients	
	REITs	Property companies	REITs	Property companies	Firm-level	Portfolio
	49,828 obs.	6,916 obs.	431 obs.			
α	0.003[a]	0.001	0.001	0.001	0.002	– 0.001
	(9.21)	(0.95)	(0.28)	(0.54)	(1.75)	(– 0.26)
β_{LIQ}	– 0.094[a]	– 0.008	– 0.095[a]	0.032	– 0.086[a]	– 0.127[a]
	(– 17.31)	(– 0.41)	(– 3.31)	(0.83)	(– 4.40)	(– 2.66)
β_{MKT}	0.678[a]	0.869[a]	0.757[a]	1.079[a]	– 0.191[a]	– 0.322[a]
	(83.19)	(30.43)	(19.00)	(20.44)	(– 6.43)	(– 4.87)
β_{SMB}	0.485[a]	0.614[a]	0.441[a]	0.404[a]	– 0.129[a]	0.037
	(47.17)	(16.82)	(8.03)	(5.55)	(– 3.39)	(0.41)
β_{HML}	0.7198[a]	0.672[a]	0.682[a]	0.471[a]	0.048	0.211[b]
	(60.83)	(16.70)	(11.45)	(5.96)	(1.11)	(2.14)
β_{UMD}	– 0.132[a]	– 0.196[a]	– 0.086[b]	– 0.309[a]	0.064[b]	0.223[a]
	(– 18.57)	(– 8.11)	(– 2.38)	(– 6.41)	(2.54)	(3.68)
Adj. R^2	19.98%	20.92%	55.66%	60.52%		

Table 5: The Liquidity Risk of REITs by Investment Focus

This table reports OLS regression estimates from a five-factor model explaining the monthly excess returns of equity REITs, by REIT investment focus. The dependent variable is the monthly return, in excess of the 1-month Treasury rate, of the CRSP/Ziman value-weighted indexes segmented by property type investment focus. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). The coefficients for these factors are β_{LIQ}, β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD} respectively. The reported *t*-statistics (in parentheses) test whether estimates are significantly different from 0. [a], [b] and [c] denote significance at the 0.01, 0.05, and 0.10 levels.

The dependent variable is the monthly excess return of the index

	Unclassified	Diversified	Health care	Industrial/ office	Lodging/ resorts	Residential	Retail	Self-storage
α	0.001	– 0.001	0.006[b]	– 0.001	– 0.003	0.002	0.002	0.005
	(0.20)	(– 0.39)	(2.42)	(– 0.18)	(– 0.91)	(0.87)	(0.97)	(1.89)
β_{LIQ}	0.021	– 0.076[b]	– 0.152[a]	– 0.122[a]	0.001	– 0.058[c]	– 0.140[a]	– 0.107[b]
	(0.50)	(– 2.23)	(– 3.70)	(– 3.22)	(0.01)	(– 1.69)	(– 3.91)	(– 2.35)
β_{MKT}	0.670[a]	0.756[a]	0.585[a]	0.808[a]	0.968[a]	0.700[a]	0.772[a]	0.525[a]
	(11.65)	(16.00)	(10.20)	(15.44)	(14.22)	(14.75)	(15.60)	(8.17)
β_{SMB}	0.506[a]	0.569[a]	0.300[a]	0.473[a]	0.734[a]	0.400[a]	0.438[a]	0.421[a]
	(6.37)	(8.72)	(3.82)	(6.55)	(7.82)	(6.10)	(6.41)	(4.78)
β_{HML}	0.409[a]	0.802[a]	0.551[a]	0.728[a]	1.049[a]	0.671[a]	0.712[a]	0.608[a]
	(4.75)	(11.33)	(6.45)	(9.29)	(10.30)	(9.44)	(9.61)	(6.33)
β_{UMD}	– 0.112[b]	– 0.095[b]	– 0.145[a]	– 0.080[c]	– 0.359[a]	– 0.018	– 0.113[b]	– 0.031
	(– 2.14)	(– 2.21)	(– 2.76)	(– 1.68)	(– 5.79)	(– 0.41)	(– 2.51)	(– 0.52)
Obs.	431	431	428	431	430	431	431	420
Adj. R^2	35.52%	50.52%	27.43%	44.85%	50.01%	42.42%	45.82%	21.06%

Table 6: The Liquidity Risk of REITs and Property Companies Across Time

This table reports OLS regression estimates from a five-factor model explaining the monthly excess returns of equity REITs and of property companies for pre-1999 and post-2003 chronological sub-periods. The table also reports the differences in estimates between the two sets of firms. In Panel A, the dependent variable is each firm's monthly stock return in excess of the 1-month Treasury rate. In Panel B, the dependent variable is the monthly value-weighted excess return of REITs and of property companies, where weights are based on each firm's past-month market capitalization. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). The coefficients for these factors are βLIQ, βMKT, βSMB, βHML, and βUMD respectively. The reported t-statistics (in parentheses) test whether estimates are significantly different from 0. a, b and c denote significance at the 0.01, 0.05, and 0.10 levels.

Panel A: The dependent variable is each stock's monthly excess return

	Pre-1999			Post-2003		
	REITs	Property companies	Diff.	REITs	Property companies	Diff.
α	0.001	<0.001	0.001	0.005[a]	0.003[c]	0.002
	(1.16)	(0.20)	(0.10)	(8.28)	(1.81)	(0.79)
β_{LIQ}	– 0.071[a]	0.068[b]	– 0.139[a]	– 0.176[a]	– 0.077[b]	– 0.099[a]
	(– 7.86)	(2.01)	(– 3.99)	(– 17.79)	(– 2.50)	(– 3.09)
β_{MKT}	0.546[a]	0.739[a]	– 0.193[a]	1.028[a]	1.094[a]	– 0.066
	(42.54)	(15.15)	(– 3.83)	(65.36)	(22.64)	(– 1.29)
β_{SMB}	0.577[a]	0.849[a]	– 0.272[a]	0.380[a]	0.736[a]	– 0.356[a]
	(30.39)	(11.46)	(– 3.55)	(14.87)	(9.05)	(– 4.18)
β_{HML}	0.551[a]	0.378[a]	0.173[c]	0.547[a]	0.268[a]	0.279[a]
	(24.23)	(4.35)	(1.92)	(21.83)	(3.47)	(3.43)
β_{UMD}	0.076[a]	0.023	0.053	– 0.223[a]	– 0.334[a]	0.111[b]
	(4.68)	(0.37)	(0.85)	(– 16.21)	(– 8.04)	(2.54)
Obs.	20,639	2,166		19,571	3,302	
Adj. R^2	12.07%	18.26%		33.10%	28.76%	

Table 6: The Liquidity Risk of REITs and Property Companies Across Time

Panel B: The dependent variable is the value-weighted portfolio monthly excess return

	Pre-1999			Post-2003		
	REITs	Property companies	Diff.	REITs	Property companies	Diff.
α	– 0.002	<0.001	– 0.002	0.004	0.002	0.002
	(– 0.98)	(0.09)	(– 0.54)	(1.06)	(0.70)	(0.33)
β_{LIQ}	– 0.051	0.116[b]	– 0.167[b]	– 0.204[a]	– 0.068	– 0.136[c]
	(– 1.52)	(1.78)	(2.28)	(– 3.35)	(– 1.25)	(– 1.67)
β_{MKT}	0.611[a]	0.948[a]	– 0.336[a]	1.098[a]	1.288[a]	– 0.190
	(14.43)	(11.39)	(– 3.60)	(11.27)	(14.89)	(– 1.46)
β_{SMB}	0.577[a]	0.534[a]	0.043	0.228	0.554[a]	– 0.325
	(8.89)	(4.19)	(0.30)	(1.39)	(3.80)	(– 1.49)
β_{HML}	0.430[a]	0.112	0.318[b]	0.594[a]	0.106	0.487[b]
	(5.99)	(0.80)	(2.01)	(3.88)	(0.78)	(2.38)
β_{UMD}	0.056	– 0.180[c]	0.236[b]	– 0.140	– 0.461[a]	0.321[a]
	(1.11)	(– 1.82)	(2.13)	(– 1.70)	(– 6.31)	(2.92)
Obs.		227			144	
Adj. R^2	59.37%	51.04%		64.89%	78.61%	

Table 7: The Liquidity Risk of REITs by Size and Dividend Frequency Portfolio

This table reports the coefficient estimates for the liquidity factor, β_{LIQ}, from OLS regressions of a five-factor model explaining the monthly excess returns of equity REITs. The dependent variable is each firm's monthly stock return in excess of the 1-month Treasury rate. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). I estimate the model for six portfolios, where each month REITs are sorted independently into three portfolios based on firm size and two portfolios based on past dividend frequency. I use the 30^{th} and 70^{th} percentiles of one-month lagged REIT market capitalizations to construct the three size portfolios. Dividend frequency is the moving average of the past eight quarters of a binary variable that equals one if the REIT paid positive dividends in a quarter, and zero otherwise. REITs with a moving average equal to one form the "=100%" dividend frequency portfolio, while REITs with a moving average less than one form the "<100%" dividend frequency portfolio. The reported t-statistics (in parentheses) test whether estimates are significantly different from 0. [a], [b] and [c] denote significance at the 0.01, 0.05, and 0.10 levels.

The dependent variable is each stock's monthly excess return

Dividend frequency portfolio			Size portfolio			
			Small	Med	Large	Diff. in β_{LIQ}
< 100%	β_{LIQ}		– 0.003	– 0.117[a]	– 0.173[a]	0.170[a]
			(– 0.12)	(– 3.91)	(– 6.49)	(4.35)
	Obs.		3,640	2,194	1,761	
	Adj. R^2		14.33%	23.56%	28.61%	
= 100%	β_{LIQ}		– 0.107[a]	– 0.102[a]	– 0.095[a]	– 0.013
			(– 8.03)	(– 10.19)	(– 10.59)	(– 0.78)
	Obs.		8,595	11,127	12,229	
	Adj. R^2		21.19%	24.97%	25.58%	
	Diff. in β_{LIQ}		0.104[a]	– 0.015	– 0.078[a]	
			(3.30)	(– 0.47)	(– 2.79)	

Table 8: The Liquidity Risk of REITs by Leverage and Mart-to-Book Portfolios

This table reports the coefficient estimates for the liquidity factor, β_{LIQ}, from OLS regressions of a five-factor model explaining the monthly excess returns of equity REITs. The dependent variable is each firm's monthly stock return in excess of the 1-month Treasury rate. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). I estimate the model for nine portfolios, where each month REITs are sorted independently into three portfolios based on firm leverage and three portfolios based on market-to-book ratios. To construct the portfolios, I use the 30th and 70th percentiles of lagged REIT leverage (total liabilities over total assets) and lagged market-to-book of assets (the sum of market capitalization and total liabilities divided by total assets). The reported t-statistics (in parentheses) test whether estimates are significantly different from 0. [a], [b] and [c] denote significance at the 0.01, 0.05, and 0.10 levels.

The dependent variable is each stock's monthly excess return

Market-to-book portfolio		Leverage portfolio: Low	Med	High	Diff. in β_{LIQ}
Low	β_{LIQ}	– 0.050[a]	– 0.125[a]	– 0.119[a]	0.069[b]
		(– 3.21)	(– 6.29)	(– 4.67)	(2.31)
	Obs.	6,151	4,894	3,714	
	Adj. R^2	17.77%	25.49%	22.29%	
Med	β_{LIQ}	– 0.115[a]	– 0.104[a]	– 0.080[a]	– 0.035
		(– 8.04)	(– 7.15)	(– 4.15)	(– 1.48)
	Obs.	4,782	5,322	5,074	
	Adj. R^2	22.21%	24.32%	19.86%	
High	β_{LIQ}	– 0.119[a]	– 0.059[a]	– 0.091[a]	– 0.028
		(– 7.72)	(– 4.21)	(– 6.38)	(– 1.35)
	Obs.	3,934	5,076	6,176	
	Adj. R^2	22.73%	19.42%	16.44%	
	Diff. in β_{LIQ}	0.069[a]	– 0.066[a]	– 0.028	
		(3.17)	(– 2.71)	(– 0.95)	

Table 9: The Liquidity Risk of REITs by Size and Operations Type Portfolios

This table reports the coefficient estimates for the liquidity factor, β_{LIQ}, from OLS regressions of a five-factor model explaining the monthly excess returns of equity REITs. The dependent variable is each firm's monthly stock return in excess of the 1-month Treasury rate. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). I estimate the model for six portfolios, where each month REITs are sorted independently into three portfolios based on firm size and two portfolios based on REIT operations type. I use the 30th and 70th percentiles of one-month lagged REIT market capitalizations to construct the three size portfolios. REITs are further classified into either development or operational firms based on their recent participation in, or avoidance of, real property development activities. The reported t-statistics (in parentheses) test whether estimates are significantly different from 0. [a], [b] and [c] denote significance at the 0.01, 0.05, and 0.10 levels.

The dependent variable is each stock's monthly excess return

REIT operations type			Size portfolio			
			Small	Med	Large	Diff. in β_{LIQ}
	Operational	β_{LIQ}	– 0.076[a]	– 0.082[a]	– 0.071[a]	– 0.05
			(– 6.17)	(– 7.59)	(– 6.20)	(– 0.30)
		Obs.	11,835	10,193	7,846	
		Adj. R^2	15.53%	21.45%	19.75%	
	Development	β_{LIQ}	– 0.179[a]	– 0.147[a]	– 0.150[a]	– 0.029
			(– 5.93)	(– 9.72)	(– 12.72)	(– 0.91)
		Obs.	2,221	5,346	7,609	
		Adj. R^2	28.04%	26.62%	30.47%	
		Diff. in β_{LIQ}	0.103[a]	0.065[a]	0.079[a]	
			(3.17)	(3.46)	(4.81)	

Table 10: The Liquidity Risk of REITs by Size and Dividend Reinvestment Plan Portfolios

This table reports the coefficient estimates for the liquidity factor, β_{LIQ}, from OLS regressions of a five-factor model explaining the monthly excess returns of equity REITs. The dependent variable is each firm's monthly stock return in excess of the 1-month Treasury rate. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). I estimate the model for six portfolios, where each month REITs are sorted independently into three portfolios based on firm size and two portfolios based on whether or not the REIT offers a dividend reinvestment plan (DRIP). I use the 30th and 70th percentiles of one-month lagged REIT market capitalization to construct the three size portfolios. The reported *t*-statistics (in parentheses) test whether estimates are significantly different from 0. [a], [b] and [c] denote significance at the 0.01, 0.05, and 0.10 levels.

The dependent variable is each stock's monthly excess return

Dividend reinvestment plan			Size portfolio			
			Small	Med	Large	Diff. in β_{LIQ}
	No	β_{LIQ}	– 0.082[a]	– 0.068[a]	– 0.116[a]	0.033
			(– 4.28)	(– 4.51)	(– 6.81)	(1.30)
		Obs.	6,301	6,113	4,391	
		Adj. R^2	15.01%	22.44%	23.04%	
	Yes	β_{LIQ}	– 0.100[a]	– 0.123[a]	– 0.108[a]	0.009
			(– 7.28)	(– 11.39)	(– 11.76)	(0.53)
		Obs.	7,755	9,426	11,064	
		Adj. R^2	20.40%	24.00%	26.45%	
		Diff. in β_{LIQ}	0.017	0.055[a]	– 0.007	
			(0.73)	(2.96)	(– 0.39)	

Table 11: The Liquidity Risk of Conditional REIT Status

This table reports OLS regression estimates from a five-factor model explaining the monthly excess returns of stocks conditional on their REIT status. I identify 440 equity REITs from the CRSP/Ziman database with return data in the Center for Research in Securities Prices (CRSP) monthly files over the interval 1980 through 2015. I exclude firm-years with a market capitalization below $20 million. For each firm-month observation I construct a binary variable (REIT status) that equals one if the monthly observation is within the first and last dates the stock is reported as having elected REIT status within the CRSP/Ziman dataset. If the monthly observation for the firm is outside of that date range, I classify the observation as not in REIT status. The dependent variable is each firm's monthly stock return in excess of the 1-month Treasury rate. As explanatory variables, the model includes the liquidity factor of Pástor and Stambaugh (2002), the market, size, and value factors of Fama and French (1993), and the momentum factor of Carhart (1997). The coefficients for these factors are β_{LIQ}, β_{MKT}, β_{SMB}, β_{HML}, and β_{UMD} respectively. I then estimate the following model:

$$R_{it} - R_{ft} = \alpha_1 Status_{it} + \alpha_0 \left(1 - Status_{it}\right) + B_1 \times Status_{it} \times F_t + B_0 \times \left(1 - Status_{it}\right) \times F_t + u_{it},$$

where F_t is a 5×1 vector containing all five factors and the B vectors contain the coefficients for the five factors. This specification allows estimation of the factor sensitivities of returns both when firms have elected REIT status (B_1) and when they have not elected REIT status (B_0). The reported t-statistics (in parentheses) test whether the coefficient estimates are significantly different from 0. [a], [b] and [c] denote significance at the 0.01, 0.05, and 0.10 levels.

The dependent variable is each stock's monthly excess return

	REIT status: Yes	REIT status: No	Diff.
α	0.003[a]	– 0.001	0.003[c]
	(9.36)	(– 0.12)	(1.68)
β_{LIQ}	– 0.096[a]	– 0.025	– 0.071[b]
	(– 17.42)	(– 0.78)	(– 2.16)
β_{MKT}	0.678[a]	0.640	0.038
	(81.99)	(12.79)	(0.75)
β_{SMB}	0.491[a]	0.274	0.217[a]
	(47.06)	(4.48)	(3.49)
β_{HML}	0.720[a]	0.641	0.080
	(60.16)	(8.62)	(1.06)
β_{UMD}	– 0.133[a]	– 0.065	– 0.068
	(– 18.55)	(– 1.38)	(– 1.43)
Obs.		49,822	
Adj. R^2		20.00%	

ESSAY 2: ON MEASURING URBAN SPRAWL

1. Introduction

The policy debate over the extent to which urban sprawl represents market failure versus how much represents efficient market outcomes has been hampered by the lack of solid empirical analysis. While conceptually straightforward, the notion that sprawl represents the dispersion of population and jobs over a wider geographic area turns out to be difficult to capture with a scalar empirical measure.[24] This paper contributes to the literature analyzing urban sprawl. It offers an empirical approach that has the potential to move the empirical study of urban sprawl to a new level.

This study examines two existing sprawl measures and proposes a new alternative adapted from financial market technical analysis. The results show that one of the older approaches is not consistent with existing land use theory. While the second older approach, the density gradient, yields results broadly consistent with land use theory, the new measure of urban sprawl offered here relaxes the stringent monotonicity constraint imposed by the traditional density gradient approach, providing results that better reflect polycentric aspects of urban areas that are also important aspects of sprawl (see Henderson and Mitra, 1996).

The debate over the extent to which urban sprawl is efficient or represents market failure has driven urban policy discussion since the start of the U.S. suburbanization process in the late 19th century (Mills and Hamilton, 1994). The neoclassical land use theory built on the seminal

[24] Notable empirical studies on polycentric city structure include McMillen and Smith (2003) and Redfearn (2007). Notable articles that develop theoretical frameworks for the polycentric city structure include Fujita and Ogawa (1982), Anas and Kim (1996), Anas and Xu (1999)

works by Mills (1967) and Muth (1969) provides a framework for understanding how real income growth, population growth, and long-run improvements in urban transportation technology drive the suburbanization of jobs and population typically identified as evidence of inefficient sprawl (Wheaton 1974, Bruckner 1987). The insights obtained from the formalization of land use theory should not be understated on this point, since the theory shows that sprawl is also the efficient market response to changes in these underlying economic factors.

In light of the insights provided by urban economic theory, the policy concerns related to urban sprawl have slowly begun to be tied to the question of how much observed sprawl is the result of efficient market responses to changes in household income, urban population, and transportation cost. Presumably, once the sprawl attributable to these factors can be quantified, the remaining extent of urban sprawl is the result of factors not included in the theory, including market failure (Mills and Lubuele 1997, Nechyba and Walsh 2004).

The seminal work by Brueckner and Fansler (1983), hereafter BF, provides widely cited empirical evidence on this point. The empirical results are consistent with urban land market theory; household income, population, agricultural land value, and urban transportation costs appear to drive the geographic extent of urban areas as predicted by theory. As important to the sprawl debate, these factors appear to explain 78% of the variation in size. Their evidence, however, is limited to small urban areas contained within single counties, a small fraction of urban areas in the U.S. in 1970 and an even less important segment of urban areas today.[25] Further,

[25] Brueckner and Fansler (1983) further constrained their sample to exclude cities with topographical irregularities, based on an implicit assumption in the Muth-Mills model that all land at any given distance from the CBD can be used for housing.

closer analysis undertaken here reveals that the original results even for that subset of urban areas are spurious; re-estimation of the BF model using repeated samples of urban areas drawn from the original 1970 Census data source yields results consistent with the original study in only 2-11% percent of samples, depending on the proxy used for transportation costs. It also turns out that the BF approach does not yield useful results for later sample periods.

Motivated by the poor performance of the BF approach, this paper turns to the venerable population density gradient (PDG) as one measure of urban sprawl. The measure is appealing in its simplicity; it depicts the percentage decline in population density with greater distance over the entire urban area. It does, however, impose monotonicity and symmetry in the population density distribution function for all directions centered on the CBD; the extent to which these characteristics distort empirical analyses is not known. While this paper uses the very popular exponential form, it should be noted that the broader literature is not in complete agreement about functional form for the population density distribution function underlying the single parameter measure. Kau and Lee (1976), for example, test linear, log, and quadratic functional forms of the population density distribution function early in the debate, and find that while the linear form can be rejected for most of the cities in their sample, the exponential form cannot. Nonetheless, they conclude that the appropriate functional form may vary across cities.

The differing degrees of polycentricity across urban areas may provide one reason why the simple density gradient approach does not work uniformly well. The growth of employment centers outside the CBD destroys the symmetry of the population distribution around the CBD as it also weakens monotonicity. The notion that the CBD is stationary over time may also be problematic. Alperovich and Deutsche (1994) find that the (functional) CBD location can vary significantly over time.

Given these potential weaknesses, this paper offers a sprawl measure based on Wilder's (1978) directional index (DX), a technical indicator of a financial security's price trend strength. While the original measure is designed to measure financial market characteristics, I borrow from the finance literature and exploit the versatility of DX by adapting it to capture urban sprawl as a dynamic process that takes place over time. It has the additional advantage of not imposing the strong monotonicity that the traditional gradient approach does. Tests of both the PDG and DX measures on a broad sample of metropolitan areas using 2000 and 2010 Census data yield stronger support for the DX measure of urban sprawl. As important to the urban sprawl policy debate, these measures of sprawl indicate only up to 34% (for PDG) or up to 40% (for DX) of the variation in sprawl can be explained by the economic factors identified by neoclassical land use theory. This is considerably less than half of the proportion found by BF. This represents a major shift in the empirical evidence underlying much of the policy debate concerning sprawl; it calls into question whether a large part of urban development experienced in the U.S. can be construed as efficient market-driven outcomes or instead represents market failure.

The rest of this paper is structured as follows. In section 2, I discuss relevant literature and important empirical and theoretical results previously obtained in the study of urban sprawl. Section 3 develops the PDG and DX measures of sprawl I use in this paper, and section 4 details my data sampling methodology and variables used. I perform empirical testing of the three approaches to measuring sprawl, test whether I can improve upon the models with additional explanatory variables, and present and discuss results in section 5. Section 6 concludes.

2. Literature Review

The Mills-Muth model depicts sprawl as an efficient market process. A tradeoff between transportation costs and land rents drives the structure of urbanized areas, with decreasing rent and population density gradients with increased distance from the CBD. In addition, the model also relies on the joint effect of rising incomes and reduced commuting costs on encouraging sprawl. The literature has thus far supported these assertions. Glaeser and Kahn (2003) show the major factor influencing sprawl is not market failure, but rather the necessity of vehicle ownership for urban living in the 20th century. LeRoy and Sonstelie (1983) model the movement of two income groups, rich and poor, into and out of city centers, and show that transportation costs are an important factor. Margo (1992) provides empirical evidence on the importance of rising incomes in driving sprawl.

Wheaton (1974) derives the following comparative statics results of the Mills-Muth model for distance to urban-rural boundary ($\bar{x}$), urban population (L), agricultural land rent (r_a), income (y), and commuting cost per round-trip mile (t), respectively,

$$\frac{\partial \bar{x}}{\partial L} > 0, \frac{\partial \bar{x}}{\partial r_a} < 0, \frac{\partial \bar{x}}{\partial y} > 0, \frac{\partial \bar{x}}{\partial t} < 0 \tag{3}$$

and BF empirically tested those predictions.[26] Specifically, BF found for a small sample of urban areas in 1970 that land area is negatively related to commuting costs ($\frac{\partial \bar{x}}{\partial t} < 0$) and agricultural land values ($\frac{\partial \bar{x}}{\partial r_a} < 0$) but positively related to incomes ($\frac{\partial \bar{x}}{\partial y} > 0$) and population ($\frac{\partial \bar{x}}{\partial L} > 0$). While their sample was heavily constrained within population bounds and that the areas be contained within a single county, their results have been widely cited as validating the Mills-Muth model.

BF used two measures of commuting costs from the Census available at the time of their study; the percent of households with at least one vehicle and the percent of commuters using public transportation, with the notions that increased vehicle ownership indicates reduced commuting costs and increased use of public transportation indicates increased (time) cost. To this end, LeRoy and Sonstelie (1983) point out that not every resident of an area commutes via the same method. In their model, the residential location choice of rich and poor groups is driven by relative income elasticities of housing and marginal commuting costs for two different methods of transportation, a faster, expensive mode and a slower, less expensive mode. In the context of BF, I consider vehicles the faster, expensive mode and public transportation the slower, less expensive mode. LeRoy and Sonstelie show that, ultimately, all residents of the poor group are able to afford the faster mode of transport as its ubiquity reduces its relative cost, which will enable this group to leave the inner city for the suburbs. The implication of this result is that as the rich move back downtown, suburban deterioration precedes urban rejuvenation.

In a more recent study, Spivey (2008) tests whether Wheaton's (1974) predictions of the Mills-Muth model hold for modern cities in the year 2000. Spivey follows BF's estimation

[26] See the appendix for derivation of the comparative statics.

methodology for *all* urbanized areas from the 2000 Census and reports results that generally support the Mills-Muth model. Moreover, she improves on the BF specification with new proxies for commuting costs and agricultural land values and additional variables to capture polycentricity and physical boundaries. However, her results are susceptible to the use of a previously estimated measure of city sub-centers for only 60 urban areas to indicate polycentricity. With a sample size of 452 urban areas, she assumes that areas for which sub-centers were not measured were monocentric. Further, the extent to which her results are driven by outliers is unknown. Unlike BF, her sample is unconstrained and includes urban areas with populations exceeding 17 million. To this point, the author notes that for a sample constrained only by the single-county requirement of BF, no stronger support for the Mills-Muth model is obtained. However, this sub-sample was not constrained by population and included urban area populations in excess of 2.7 million.

At issue with the prior empirical studies is whether Wheaton's (1974) result that the scalar measure of urban spatial size is sufficient to capture the process of urban sprawl. Indeed, sprawl is usually defined as an expansion of a city footprint accompanied by a lower urban population density. Edge cities are a prime example of sprawl. Typically characterized by low-density suburbanization and employment (Glaeser and Kahn, 2003), edge cities derive their name from their location on the urban fringe, which leads to loss in welfare resultant of the loss in open space. However, Nechyba and Walsh (2004) point out that while most monocentric city models constrain the definition of open space to rural areas on the urban fringe, there is no reason to believe that households highly value such open space, and that open space within a suburb is valued more than that at the fringe. The literature supports this claim, with empirical findings showing a nonlinear relationship between house prices and open space proximity.[27]

Nonetheless, the size of an area doesn't account for *how* people become dispersed throughout an urban area over time, and the literature has thus far struggled to capture the time-series nature of sprawl from cross-sectional data. On one hand, the PDG improves over the land area measure by indicating the degree to which residents are moving away from the city center. Yet, the PDG relies on strict monotonicity of the population distribution function, which may or may not hold. On the other hand, the land area of an urban area fails to capture any information about the density gradient and can be heavily influenced by the existence of an edge city or a few houses popping up on previously agricultural land. In addition, both the PDG and spatial size implicitly assume that the momentum of past population changes can be inferred from cross-sectional data.

I consider two examples to help illustrate the above short-comings of the PDG and land area as measures of sprawl. First, it may be that the size of a closed city expands as a few residents move outward from the CBD to previously agricultural land, while the majority of residents live in denser housing relatively closer to the city center due to employment opportunities. This may be due to positive income and population shocks, as income variations across residents contribute to the spatial distribution of housing location choice. Second, it may be the development of a bustling planned community on the urban fringe that is causing the foot print of an area to expand, and the spatial expansion is accompanied by a sizeable portion of the population moving to the

[27] See Geoghegan et al. (1997).

new community. The former scenario is consistent with Wheaton (1974), who shows that for a monocentric urban area, an increase in population can lead to higher structural densities as the producers of housing substitute away from land due to higher land prices, while at the same time the increase in population causes an urban area to expand. The latter case is just as likely, yet the PDG and spatial size metrics fail to differentiate between the two scenarios. The measure developed herein deals with the inability of prior measures to jointly capture the spatial and population changes within the area and aims to capture the strength of population movement trends, not simply land area expansions.

3. Constructing a Measure of Sprawl

3.1 Population Density Gradient

The PDG used in this paper begins with calculating Census tract-level population densities as

$$Density_{ji} = \frac{Population_i}{Area\ in\ mi_i^2} \quad (4)$$

where the subscripts i index each tract and j each urban area. Following the findings of Kau and Lee (1976), the PDG is obtained as β from the following regression for each urban area

$$lnDensity_{ji} = \alpha + \beta Distance_i + \varepsilon_i \quad (5)$$

where $Distance_i$ is the average distance, computed using the Vincenty geodesic, between each CBD centroid and the centroid of the i^{th} tract.

3.2 Directional Index for Securities Prices

Wilder (1978) developed directional indicators for the technical analysis of commodities price movements, which have since been applied to almost all securities prices. Though Wilder (1978) developed these indicators across time and price-space, I adapt them for use across time and spatial deciles based on distance from the CBD. I begin by developing the directional index used for the technical analysis of securities prices. With notation borrowed from Lam and Chong (2006), the first in a series of intermediate calculations are $+DM_t$ to indicate positive price movement and $-DM_t$ to indicate negative price movement, where for time t, PH_t is the intraday high and PL_t is the intraday low security price.

$$+DM_t = \begin{cases} PH_t - PH_{t-1} & \begin{Bmatrix} PH_t - PH_{t-1} > 0 \; and \\ PH_t - PH_{t-1} > PL_{t-1} - PL_t \end{Bmatrix} \\ 0 & Otherwise \end{cases} \tag{6}$$

$$-DM_t = \begin{cases} PL_{t-1} - PL_t & \begin{Bmatrix} PL_{t-1} - PL_t > 0 \; and \\ PL_{t-1} - PL_t > PH_t - PH_{t-1} \end{Bmatrix} \\ 0 & Otherwise \end{cases} \tag{7}$$

Note that for any given t, both $+DM$ and $–DM$ must both be non-negative and at least one must be positive. For each t, a single value is obtained that indicates whether 1) for $+DM > 0$, the intraday high has moved higher from its prior intraday high than its low has moved lower from its prior intraday low or 2) for $–DM > 0$, the intraday low has moved lower from its prior intraday low than its high has moved higher from its prior intraday high.

Next, the true range, TR_t, is defined as

$$TR_t = max\{PH_t - PL_t,\ PH_t - P_{t-1},\ P_{t-1} - PL_t\}, \tag{8}$$

where P_t is the security closing price at time t. As the true range measures the trading range of a security price, it proxies for volatility by taking the maximum of 1) the difference between the current intraday high and low prices, 2) the difference between the current intraday high and the prior day's closing prices, and 3) the difference between the prior day's closing and current intraday low prices. The true range and directional movement indicators are then combined to form N-day directional indicators, $+DI_N(t)$ and $-DI_N(t)$, defined as

$$+DI_N(t) = \frac{\sum_{i=t-N+1}^{t} +DM_i}{\sum_{t-N+1}^{t} TR_i} \tag{9}$$

$$-DI_N(t) = \frac{\sum_{i=t-N+1}^{t} -DM_i}{\sum_{t-N+1}^{t} TR_i} \tag{10}$$

where i indicates the time reference and N the number of days over which the directional movement indicators are summed. $+DI_N$ and $-DI_N$ convert the directional movement of a security price into a proportion of that security's true range, with higher values of $+DI_N$, relative to $-DI_N$, implying relatively smaller downward price movements and higher values of $-DI_N$, relative to $+DI_N$, implying relatively smaller upward price movements.

By taking the difference between the N-day directional *indicators* as a ratio to their sum, the N-day directional *index* combines the information in Equations 9 and 10 to compute a daily indicator of the strength of the price trend, and is valued between 0 and 100, with higher values indicating a stronger price trend. The directional index is calculated as

$$DX_N(t) = \left|\frac{(+DI_N(t)) - (-DI_N(t))}{(+DI_N(t)) + (-DI_N(t))}\right| * 100 \tag{11}$$

Wilder (1978) continues to develop Equation 11 into the average directional movement index by taking its simple moving average, which indicates the relative strength of the price movement over a given time period. However, as the index developed in this paper will be considered over two time periods, development beyond Equation 11 is not necessary.

3.3 Adapting the Directional Index to Measure Sprawl

The first step in adapting Equation 11 to measure sprawl is to divide the urban areas into deciles, denoted $d_i, i \in [1, 10]$, based on the spatial size in square miles of the areas. I determine decile placement of each tract in a given area by using the Vicenty geodesic to calculate the distance from the CBD centroids to each tract centroid. For each urban area, the decile populations are then computed with tract-level data. In doing so, I am able to attain the granularity of frequent financial transactions needed to adapt *DX* to measure sprawl. The choice of deciles, rather than a finer or broader division, is motivated by the need to achieve information about changes occurring in the urban area, while at the same time mitigating uninformative computational excess. After computing each decile-population, the directional movement indicators of Equations 6 and 7 are calculated, for urban area j, decile i, and time *t* as

$$+DM_{j,i} = \begin{cases} p_{i,t} - p_{i,t-1} & \{p_{i,t} - p_{i,t-1} > 0\} \\ 0 & Otherwise \end{cases} \tag{12}$$

$$-DM_{j,i} = \begin{cases} p_{i,t-1} - p_{i,t} & \{p_{i,t-1} - p_{i,t} > 0\} \\ 0 & Otherwise \end{cases} \tag{13}$$

where $DM_{j,i}$ is the directional movement, either positive or negative, and $p_{i,t}$ is the population for decile *i*. In this adaptation, $+DM_{j,i} > 0$ indicates population growth in decile *i* for time period *t,* while $-DM_{j,i} > 0$ indicates that the population has decreased in that decile and time period. The true range of Equation 8 is then calculated as

$$TR_j = max(+DM_{j,i}, -DM_{j,i}), \quad (14)$$

and, instead of a volatility proxy as in Equation 8,Equation 14 provides information on the maximum decile population change of urban area j for later use in determining proportional changes. Once TR is obtained, the directional indicators defined in Equations 9 and 10 are computed below as Equations 15 and 16, except now the numerator is summed across space (index i is a decile, instead of a time, reference) and the denominator is not a sum, as the true range is now computed over two time periods.

$$+DI_j = \frac{\sum_{i=1}^{10}(+DM_{j,i} * w_i)}{TR} \quad (15)$$

$$-DI_j = \frac{\sum_{i=1}^{10}(-DM_{j,i} * w_i)}{TR}. \quad (16)$$

A notable change in Equations 15 and 16 from Equations 9 and 10 is that a weighting system must be employed in order to develop a meaningful measure of sprawl, or arbitrary values are obtained. I use a distance-weighted k-nearest-neighbor rule (Dudani, 1976), where $w_1 = 1$, $w_i = \frac{dist_{10} - dist_i}{dist_{10} - dist_1}, i \in [2, 9]$, and $w_{10} = \frac{1}{dist_{10} - dist_1}$, where subscripts denote the decile, and $dist_i$ is the average distance, for each urban area, from the center of the CBD to the centroids of the tracts contained within corresponding decile, i. This is a naturally intuitive choice for the weighting rule as it is non-probabilistic, assigns varying weights for each urban area based on distance from the CBD to urban boundary, and assigns decreasing weights for each decile with distance from the CBD. As with the PDG, the Vincenty geodesic is used to compute the CBD centroid-to-tract centroid distances. The importance of the directional indicators is that each one aggregates all of the information from the weighted decile population changes as a ratio to the maximum decile population change for a given urban area, with $+DI$ and $-DI$ aggregating those deciles for which population increased or decreased, respectively. Finally, the directional index, which captures the direction and magnitude of the population density gradient, is defined as

$$DX_j = \frac{(+DI_j)-(-DI_j)}{(+DI_j)+(-DI_j)} * 100. \tag{17}$$

3.4 Empirical Characteristics of the Directional Index

One important change in Equation 17 from Equation 11 is that the absolute value of the quotient is not used, so DX is now bounded by $[-100, 100]$ and the measure can assume negative or positive values. A negative value indicates that the population density gradient is shallower, suggesting sprawl. For instance, consider Pocatello, ID and Tuscaloosa, AL, for which directional index values of -58.21 and -29.70 are obtained, respectively. Having negative values indicates that both area populations are expanding toward their peripheries, while Pocatello has a stronger trend of sprawl over the measurement period than does Tuscaloosa. Conversely, a positive value for the directional index indicates that the population density gradient is steeper, or that there is relatively more population growth towards the CBD. Again, consider two areas, Danbury, CT and Missoula, MT, with DX values of 44.62 and 18.95, respectively. In this example, Danbury experienced more population growth toward the CBD than Missoula, MT over the measurement period.

Further, it is worth noting a few important characteristics of *DX* that affect its interpretation and implementation. First, while the directional index benefits from relaxing the strict monotonicity of the *PDG* and is thus more informative with respect to asymmetric growth and polycentric urban structure, it is unable to indicate *which* decile is experiencing the population change. For example, *DX* cannot explicitly differentiate between population growth in decile 10, at the extreme periphery, and growth in deciles 8 or 9. However, its strength lies in capturing population movement *trends* either toward or away from the CBD. Second, as it measures population changes, *DX* can take on a negative value, indicating sprawl, for an area that spatially contracted but experienced relative population growth in the outer deciles. Conversely, *DX* can take on a positive value, indicating urban contraction, for an area that spatially expanded but experienced relative population growth toward the CBD. Consider population and land area changes for the urban areas discussed above, which are presented in Table 12. In panel A of Table 12, Pocatello and Tuscaloosa experienced both population growth in the outer deciles toward their urban peripheries and contraction in the inner deciles closer to their CBDs, yet Tuscaloosa spatially contracted from 2000 to 2010. This is evident by the lower directional index value for Tuscaloosa, suggesting weaker sprawl. In panel B of Table 12, Danbury and Missoula spatially expanded while their populations declined in the outermost deciles, and Danbury experienced much stronger population growth toward the CBD than Missoula. Lastly, a value of zero for *DX* indicates that the population has remained constant in each decile. Empirically, however, I did not have any such cases of a stagnant population in any of the samples.

4. Data

Testing on whether the BF methodology lead to spurious results is performed using data from the 1970 decennial U.S. Census and is collected from the 1972 County and City Databook. Specifically, the variables I collect are median income ($MedInc$), the percent of households with at least one vehicle ($Vehicles$), the percent of residents commuting via public transit, excluding taxis ($Transit$), land area in square miles ($Area$), and population (Pop). Agricultural land rents ($AgValue$) are collected from the 1969 USDA Census.[28] To empirically test the PDG and DX, I begin with the population of all urban areas as defined by the U.S. Census Bureau for the years 2000 and 2010. Eliminating observations for which a missing value prevented complete calculations of either the PDG or the DX, resulted in 440 and 308 observations for the years 2000 and 2010, respectively. In addition to those variables noted above, I also collect the location coordinates for the centroid of each urban area from the Census bureau.[29] For the year 2000, data are obtained from the decennial Census. However, as the "short form" Census was conducted for the 2010, American Community Survey 5-year estimates are used for 2010 data, with the exception of the urban area population, which is obtained from the 2010 decennial Census. Finally, the rental value of agricultural land is collected from the USDA Census for the years 2002 and 2012. Lists of the Census tracts contained within each urban area, with tract centroid coordinates, population, and containing county, are obtained for the years 2000 and 2010 from the Missouri Census Data

[28] The USDA Census is not conducted decennially, and the years for which it is conducted do not coincide with the U.S. Census Bureau decennial program

[29] The urban area coordinates obtained from the U.S. Census Bureau are not necessarily for the precise urban area centroid, but are a close approximation, according to a Census Bureau employee.

Center, with centroid coordinates for tracts partially contained within an urban area weighted by the proportion of the population of the tract that resided within the urban area.

BF sample 40 urban areas from the 1970 U.S. Census. Their relatively small sample size is dictated by the constraints that the urban areas are contained within single counties and have populations inclusively bounded by 52,000 and 257,000. The motivation for these constraints is to "accurately measure the value of agricultural land immediately adjacent to the built-up part of the city" (Brueckner and Fansler (1983, p. 481)). However, as Nechyba and Walsh (2004) argue that the Census defined urban areas accurately represent built-up urban and suburban Census blocks, I test whether the constraint imposed by BF is necessary and leads to spurious results in two ways: 1) sample the population of urban areas from the 1970 Census and, after applying the population and single-county constraints, obtain 79 observations, from which I create 1,000 subsamples of 40 urban areas and 2) out-of-sample tests from the 2000 and 2010 Censuses both with and without the population and single-county constraints. After applying the sample constraints noted above, the constrained sample has 48 and 103 urban areas for 2000 and 2010, respectively.

Additionally, I test alternate measures of commuting costs obtained from the Texas A&M University's (TAM) Transportation Institute for the years 2000 and 2010. The institute compiles an Urban Mobility Scorecard annually dating back to 1982 for 101 urban areas.[30] Due to the fact that these commuting cost measures are available for limited number of urban areas, I only

[30] For a detailed discussion of the methodology used to compute the various measures obtained from the Urban Mobility Scorecard, see Appendix A; Methodology for the 2015 Urban Mobility Scorecard, accessible via http://mobility.tamu.edu/ums/congestion-data/

compiled unconstrained samples for 2000 and 2010, comprising 97 and 83 urban areas, respectively. The loss in observations from 101 urban areas is due different naming conventions between the U.S. Census and TAM urban areas, which I was unable to match with certainty. The variables of interest are freeway (*Freeway*) and arterial street *(Arterial*) daily vehicle miles of travel per auto commuter, gallons of excess fuel consumption due to travel delay per auto commuter (*Gallons*), annual hours of delay per auto commuter (*Delay*), travel time index (*TTI*), stress index (*Stress*), and congestion costs (*Congestion*). Freeway and arterial miles driven, excess fuel consumption due to delay, or annual hours of delay are self-explanatory. The travel time index is simply a measure to capture the ratio of delayed travel time during peak travel hours to non-delayed travel time. While the travel time index includes travel in all directions during the peak periods, the stress index is computed similarly, except it considers only travel in the peak direction during the peak periods. In essence, the stress index is intended to capture the daily delay of commuting to work. Further, congestion costs are measured in dollars, and are estimates of the costs of excess fuel consumption and time delay.

Contrary to notion of rising incomes encouraging sprawl, Turnbull (1998) shows that multiple worker households and job site uncertainty, whether from greater job mobility or employment insecurity coupled with multiple potential job sites in the urban area, pull households to more central residential locations than predicted for single worker households under jub site certainty. Intuitively, centrally located residential sites generate both lower expected commuting costs and variance in commuting costs when there is a probability that both workers will not be employed at the same urban location. This result suggests that I should include a proxy for dual income households in the empirical models. I include the percentage of the civilian workforce that is female (*Female*) for each urban area. This variable is drawn from the U.S. Census.

Descriptive statistics for the year 1970 data are presented in Table 13. Taking a random sample of 40 urban areas from these data, as in BF, may very well lead to spurious results, due to outliers with respect to the unconstrained variables. Descriptive statistics for the years 2000 and 2010, for the unconstrained and constrained samples, are reported in Table 14 and Table 15, respectively. These samples will be used to test the *PDG*, as well as, whether the BF methodology is robust to out-of-sample tests.[31] For both 2000 and 2010, the *PDGs* have maximum values that are positive, except for the constrained 2000 sample. I investigate the sample further and perform regression analysis both with and without those urban areas for which the *PDG* was positive and obtain results that are qualitatively identical and quantitatively similar, with only a negligible improvement in adjusted R^2s. The results obtained without positive-gradient urban areas are not reported but are available from the author upon request.

Turning to Table 14 and Table 15, there are considerable ranges for the variables, attesting to the fact that urban areas are substantially heterogeneous. For example, considering the average value of agricultural land, it can be inferred from the minimum and maximum values that the premium agricultural land carries in some locales constrains sprawl more so than it does in others, a point consistent with Wheaton (1974). Descriptive statistics for the directional index are displayed in Table 16 for the unconstrained (panel A) and constrained (panel B) samples. For the unconstrained sample, the average urban area has a *DX* value of 23.24, suggesting that the urban

[31] Chow tests using both commuting cost proxies rejected the null hypothesis of pooling the 2000 and 2010 samples at the 10% level or better.

areas were contracting between 2000 and 2010, while for the constrained sample, the average value of -70.17 suggests that the urban areas were expanding. The average directional index values for the two samples indicate that the sample constraints imposed by BF lead to the selection of urban areas that better fit the theory, an assertion that is tested in the next section.

5. Empirical Analysis

5.1 Subsample Analysis

BF find that the land area of an urban area is positively related to median income and population, while it is negatively related to the value of agricultural land and commuting costs. They measure commuting costs using either the percent of households with at least one vehicle or the percent of commuters using public transit, excluding taxis. While an increase in the usage of public transit reflects an increase in commuting (time) costs, an increase in vehicle ownership represents a decrease. I perform testing by estimating Equation 18 on 1,000 repeated samples of 40 observations drawn from the 79 urban areas that fall within the constraints noted in section 4. Though BF perform regressions using both OLS and the Box-Cox (1964) transformation, they note that the results are qualitatively similar. For comparability across results in this paper, I only estimate the following linear specification

$$Area_j = \propto + \beta_1 Pop_j + \beta_2 MedInc_j + \beta_3 CommutingCosts_j + \beta_4 AgValue_j + \varepsilon_j, \qquad (18)$$

where the coefficient expectations are $\beta_1 > 0$, $\beta_2 > 0$, $\beta_4 < 0$, and $\beta_3 > 0$ if vehicle ownership is used (model I), or $\beta_3 < 0$ if public transit use is used (model II) to proxy for commuting costs. Table 17 reports hypothesis tests of the results from estimating Equation 18. Hypotheses, where the null is in agreement with the BF findings, cannot be rejected for all variables, except median income, at the 1% level. However, I find that the coefficient on the value of agricultural land is positive in about 20% and 8% of regressions for models I and II, respectively. Further, though I reject the null for median income, I find that it is positive in only 10.8% of the regressions for model II. Comparing this to the quantity for which the coefficient on median income is positive in model I (2%), it appears that public transit usage is a better proxy for commuting costs than vehicle ownership.

The results of subsample testing suggest that the BF results are highly sample dependent. This is further illustrated by the distributions of coefficients from estimating Equation 18 with 1,000 repeated samples in Figure 2 through Figure 5 for model I and Figure 6 through Figure 9 for model II, followed by descriptive statistics of the coefficients for both models in Table 18 (panel A) and linear regression results reproduced from BF (panel B). The distributions in Figure 2 through Figure 9 visually illustrate the vast range of coefficient estimates and further evince how highly sample dependent the results obtained by BF are. For instance, consider median income. BF obtained a coefficient estimate of 0.00624 on *MedInc* and (panel B,Table 18) for the model using *Vehicles* as the commuting cost. Compared to the average coefficient estimate from repeated samples in Table 18 (panel A) for *MedInc* of $-3.15*10^{-3}$ for the comparable model, BF's estimates are remarkably different. Indeed, I find that the mean coefficient on median income in model I is over two standard deviations from being positive and over six standard deviations away from BFs estimate. Continuing with Table 18, the magnitudes of the mean coefficients from repeatedly sampling the 1970 data (panel A) are strikingly different from those obtained by BF (panel B). Most notably is the mean coefficient on median income, which is negative in the repeated samples, but positive in the BF results. Moreover, the only coefficients that are quantitatively similar to those found by BF are those on *Pop*, yet the current estimates are about one hundred times smaller.

5.2 Out-of-sample Tests

In light of the results presented in the previous section, I perform out-of-sample testing of the model in Equation 18 with 2000 and 2010 Census samples for the same variables, the results of which are reported in Table 19. Overall, the results from out-of-sample tests do not align with the conclusions of BF, and the model appears to explain about 9-10% of the cross-sectional variation in urban land areas. Median income is positive only in the unconstrained sample for year 2000 and significant only when using public transit usage as a proxy for commuting costs, while it is negative in all other regressions and significantly so for the constrained sample. As with the results from subsample testing, population is positive and significant across both samples and years. The value of agricultural land is unexpectedly positive and significant for all regressions, except the year 2000 constrained sample, where the coefficients are not significant. The coefficients on vehicle ownership are of the expected positive signs, though none is significant in any models, and public transit usage is of the expected negative sign in most of the models, and significant in only the year 2010. Thus, the results from out-of-sample tests of the BF model evince the highly sample-dependent nature of their results, with the only variable to reliably support the BF findings in out-of-sample tests being population.

5.3 Population Density Gradient

In that the BF results cannot be replicated in- or out-of-sample, I turn to the PDG as another measure of urban sprawl. While the PDG is appealing in its simplicity and ease of interpretation, as noted in section 1, it suffers from imposing strict monotonicity and symmetry constraints on the population density distribution function. Nonetheless, in light of the failure of the BF model shown earlier, I compare the *PDG* against the urban area size by regressing the *PDG* on the unconstrained and constrained samples for 2000 and 2010, as follows

$$PDG_{j,t} = \propto + \beta_1 Pop_{j,t} + \beta_2 MedInc_{j,t} + \beta_3 CommutingCosts_{j,t} + \beta_4 AgValue_{j,t} + \varepsilon_{j,t} \quad (19)$$

the results of which are presented in Table 20. A negative coefficient in Table 20 suggests that a variable encourages urban sprawl, as the *PDG* becomes flatter. Conversely, a positive coefficient indicates a steeper PDG and thus diminishes sprawl. The coefficient expectations are $\beta_1 < 0$, $\beta_2 < 0$, $\beta_4 > 0$, and $\beta_3 < 0$ if vehicle ownership is used, or $\beta_3 > 0$ if public transit use is used. Although median income, population, and public transit usage are all of the expected signs, median income is significant in fewer than half of the regressions, and transit usage is significant in only one. The negative coefficient on the value of agricultural land in both unconstrained models for year 2000 and the unconstrained model for year 2010 when *Vehicles* is used is unexpected and puzzling. As expected, the coefficient on *AgValue* is positive in all other regressions, but is not statistically significant. Though public transit usage may be a better proxy for commuting costs than vehicle ownership, as it has positive coefficients in all regressions, the coefficient is only significant in the year 2000 unconstrained sampled.

Even though the results in Table 20 are generated using the *PDG*, the use of both unconstrained and constrained samples is intended to test, not only an alternate measure of sprawl, but also the sampling methodology of BF. However, for the constrained sample, the model explains about 6% of the cross-sectional variation in *PDG* for 2010, while it is able to explain approximately 27% of the variation in PDG for 2000. This result is anomalous, as the model consistently explains approximately 7-8% of the variation in PDG for both years in the unconstrained sample. However, the results appear to be improved by constraining the sample for the year 2000.

5.4 Directional Index

In light of the mediocre performance of the PDG in the previous section, I now test the directional index, which relaxes the monotonicity imposed by the PDG. To test the directional index, the following model is estimated

$$DX_{j,t} = \alpha + \beta_1 \Delta MedIncome_{j,t} + \beta_2 \Delta Commuting\ Costs_{j,t} + \beta_3 \Delta AgValue_{j,t} + \beta_4 Control_{j,t} + \varepsilon_{j,t} \quad (20)$$

where the changes in the level variables are used as independent variables. $\Delta Commuting\ Costs$ are either the point-change in the percent of households that own at least one vehicle ($\Delta Vehicles$) or the point-change in the percent of commuters using public transit, excluding taxies, ($\Delta Transit$) and $Control$ is either the year 2000 land area in square miles ($Area$) or the year 2000 population (*Pop)*. The results of estimating Equation 20 for the unconstrained and constrained samples are reported in Table 21.

I expect $\beta_1 < 0$, $\beta_3 > 0, \beta_4 < 0$, and $\beta_2 < 0$ if vehicle ownership is used, or $\beta_3 > 0$ if public transit use is used. Overall, the results are improved over the land area or PDG measures, especially for the constrained sample, and *Transit* continues to a better proxy for commuting costs than *Vehicles*. The change in median income is statistically significant in all regressions, but opposite from the expected sign for the unconstrained sample. Vehicle ownership is of the expected negative sign in three out of four regressions yet significant in only one. On the other hand, public transit usage is positive and statistically significant in all regressions, which, like the results in section 5.1 Subsample Analysis, suggest that it is a better proxy for commuting costs than vehicle ownership. The value of agricultural land, 2000 population, and 2000 land area are all significant at the 1% level and of the expected signs. Additionally, the adjusted R^2 values of 27-40% for the constrained sample vs. 16-21% for the unconstrained sample further suggest that constraining the sample may lead to better results. Thus, it appears that DX is a better measure of sprawl than land area or the PDG.

5.5 Alternate Measures of Commuting Costs

I now consider other measures of commuting costs. Perhaps, the factors chosen by BF are indeed the underlying economic forces that drive the process of sprawl, but the proxies for commuting costs can be improved. In this section, I test measures of TAM commuting costs discussed in section 4. Data. These measures are all expected to have positive coefficients, as an increase in any one of these variables represents an increase in commuting costs and should decrease sprawl. As the directional index inherently measures change, I compute the changes in the variables for use in regressions of DX. The change variables are preceded by a Δ in the relevant output tables. An important note should be made on the interpretation of the changes the travel time and stress indices. For example, as TTI is a ratio of peak travel time to non-peak travel time, a value of 1.11 the year 2000 indicates that it would take an auto commuter 11% longer during peak hours than non-peak hours to complete the same trip. If in the year 2010 the TTI for the same urban area were 1.22, one might be inclined to think that this represents an increase of 11 percentage points, but the reader would be mistaken. An increase from 1.11 to 1.22 represents a 100% increase in TTI, as it is a doubling from 11% to 22% in travel times during peak hours over non-peak hours. $Stress$ is interpreted similarly.

<u>5.5.1 Population Density Gradient</u>

The results of estimating Equation 19 with the TAM measures of commuting costs are reported in Table 22 (2000) and Table 23 (2010). Though the signs on the commuting costs coefficients are mixed, none of the coefficients is significant for 2010, and only $Delay$ and $Congestion$ are for 2000. However, the signs on $Delay$ and $Congestion$ are opposite from those expected. The negative coefficient on $MedInc$ is as expected in both years, yet it is not statistically significant in any model in 2010, while it is significant in every model in 2000, except when TTI is used. That population is negative is as expected, though it is not significant in regressions where $Gallons, Delay, TTI,$ or $Congestion$ is used with the year 2000 samples. Further, while the negative coefficient on $AgValue$ is not as expected, it is not significant in any regression.

<u>5.5.2 Directional Index</u>

Table 24 and Table 25 present the results of estimating Equation 20 with the TAM measures of commuting costs and either Pop or $Area$ as the control, respectively. Most notably is the opposite and statistically significant income effect from that expected, which is puzzling, especially considering that both population and area are both negative and significant. $AgValue$ is of the correct positive sign when $Freeway$, $Arterial$, or $Gallons$ is used with $Area$, but only when $Freeway$ or $Arterial$ is used with Pop. However, the value of agricultural land is not significant in any regression. Interestingly, when population is used, all of the alternate commuting cost measures are significant and opposite in signs from those expected, but when land area is used, only *gallons* is of the correct positive sign and significant. The TAM commuting cost measures do not seem to perform better than public transit usage or vehicles in either the PDG or DX regressions. I further examine these measures in section, after testing whether dual-income households affect residential location choice.

5.6 Dual-Income Households

In the previous section, I tested various alternate commuting cost measures, and none emerged as being empirically reliable. Now, I examine whether dual-income households are indicative of minimizing transportation costs by living centrally, and thus encouraging urban contraction (Turnbull, 1998). I chose to use the percent of the civilian labor force that is female ($Female$) as a proxy for dual-income households.

5.6.1 Population Density Gradient

I augment Equation 19 with $Female$ as

$$PDG_j = \propto + \beta_1 Pop_{j,t} + \beta_2 MedInc_{j,t} + \beta_3 CommutingCosts_{j,t} + \beta_4 AgValue_{j,t} + \beta_5 Female_{j,t} + \varepsilon_{j,t}. \quad (21)$$

The results of estimating Equation 21 are presented in Table 26. If dual-income households cause sprawl to diminish, then $\beta_5 > 0$. Again, the model fits the 2000 constrained sample remarkably better than any other sample, with adjusted R^2 values of 49%, compared to 9-11% for the unconstrained and 2010 constrained samples. The coefficient of interest here is β_5, which is positive and significant in all year 2000 samples. Although β_5 is mixed in sign for 2010, it is not significant in any model. Of the other variables, the coefficients on income and population are reliably negative, but income is not significant for the year 2010, while population remains significant across samples and years. $AgValue$ remains statistically insignificant, as do $Vehicles$ and $Transit$.

5.6.2 Directional Index

By adding $\Delta Female$ to (14), I test the effect of dual-income households with the directional index as

$$DX_{j,t} = \propto + \beta_1 Pop_{j,t} + \beta_2 MedInc_{j,t} + \beta_3 CommutingCosts_{j,t} + \beta_4 AgValue_{j,t} + \beta_5 Female_{j,t} + \varepsilon_{j,t} \quad (22)$$

where $\Delta Female$ is the point-change in female participation in the civilian labor force from 2000 to 2010. Results from estimating Equation 22 are reported in Table 27. I expect $\beta_5 > 0$. The first thing to notice is that the adjusted R^2 values remain effectively unchanged for the unconstrained samples but are substantially reduced for the constrained samples. This is may be due transportation costs not being *as* much of a factor in housing location choice for dual-income households residing in a small, single-county urban areas, as these costs may be for a similar household that resides within a large urban area. However, the coefficient on $Female$ is unexpectedly negative, though significant only in the year 2010 constrained samples. Also unexpectedly, $MedInc$ and $Vehicles$ are positive, but only the coefficient on income is significant in the unconstrained sample and one constrained sample regression. $Transit$, Pop, and $Area$ are all of the expected signs and significant. $Female$ as a proxy for dual-income households does not perform empirically as expected.

5.7 Dual-income Households with Alternate Measures of Commuting Costs

Because I obtained mixed results in the prior two sections, in this section, I combine female labor force participation and the TAM commuting costs.

5.7.1 Population Density Gradient

I re-estimate Equation 21 with the TAM measures of commuting costs, the results of which are in Table 28 (2000) and Table 29 (2010). Overall, the model for the year 2000 works better than for the year 2010, a theme that has been somewhat consistent and puzzling throughout this study with the PDG. In Table 28, $MedInc$, Pop, $AgValue$, and $Female$ are all of the expected signs, though only $AgValue$ is not significant. Further, $Delay$ and $Congestion$ are both negative and significant, suggesting that they contribute the sprawl, opposite from what I expected. Turning to Table 29, only Pop is significant. The adjusted R^2 values of 23-28% for year 2000 and 32% for all regressions for year 2010 are improved over the results with either $Female$ or the TAM commuting costs.

5.7.2 Directional Index

The results of re-estimating Equation 22 with the TAM commuting cost measures are reported in Table 30 and Table 31 with population and area as the control variables, respectively. The results are mixed, with $Area$ emerging as a better control variable than Pop as measured by the R^2 values. This finding in regressions of DX has been consistent throughout this paper. Another finding throughout regressions of DX is the opposite income effect from that expected. In both Table 30 and Table 31, income is positive and significant in all models, a result that remains perplexing. Further, *AgValue* is not significant in any model, though it is correctly positive in all regressions where *Area* is the control, and *Pop* and *Area* remain negative and significant. As for the TAM commuting costs, in the *Area* regressions, *Freeway* and *Arterial* are unexpectedly negative and significant. Yet in the *Pop* regressions, every TAM measure is significant and negative. It seems as though *Transit* is the only empirically reliable measure of commuting costs to emerge. Consistently throughout the tests in the paper, it is the one measure to withstand different measures of sprawl, and different model specifications, while remaining positive and significant.

5.8 The Influence of Very Large Urban Areas

The review of the U.S. metropolitan growth experience in the 1990s by Glaeser and Shapiro (2003) suggests that, for a variety of reasons, the largest metropolitan areas in the U.S. exhibit spatial distributions of growth that differ from most smaller urban areas. This section investigates how the gradient and *DX* sprawl measures perform when the five largest urban areas are omitted from the sample. For brevity, I report only results from estimating Equations 21 and 22 in the two sections that follow. Results from estimating Equations 19 and 20 with the TAM commuting cost measures, without the female labor force participation variable, are qualitatively similar and can be obtained from the author.

5.8.1 Population Density Gradient

Results of re-estimating Equation 21, less the five largest urban areas by population, are reported in Table 32 and Table 33 for the years 2000 and 2010, respectively. *MedInc, Pop*, *Transit*, and *Female* are all of the expected signs and significant for 2000. The same is true for 2010, except *Transit* and *Female* are not significant. Additionally, none of the TAM commuting costs are statistically significant.

5.8.2 Directional Index

The results presented in Table 34 and Table 35 are from re-estimating Equation 22, less the five largest urban areas by population. The positive coefficient on *MedInc* is significant and has persisted throughout testing in this article. This result is remains unexpected and puzzling. *AgValue* is of the expected positive sign and significant in the models using *Vehicles* or *Transit*. Though the sign on *AgValue* is mixed in the models using the TAM commuting costs, none of the coefficients on *AgValue* is significant, while all of the coefficients on the TAM commuting costs are opposite from that expected and significant. Further, *Female* is unexpectedly negative and significant in every regression, except the model with *Area* as the control and *Gallons* as the commuting cost, where it is not significant.

The results presented in the preceding sections point to a several conclusions. First, the results reported by BF are highly sample dependent. Their results cannot be replicated and are robust to neither repeated in-sample nor out-of-sample tests. Second, the directional index developed herein appears to better capture the non-monotonic process of sprawl than the population density gradient. The results obtained for the *PDG* suggest a break in the data between the years 2000 and 2010. Indeed, I rejected pooling of the 2000 and 2010 samples based on the Chow test noted above. The *DX* does not indicate any such break, and is thus more flexible and better able to handle a wider variety of urban configurations than *PDG*. One result that is consistently puzzling is the significant and positive income effect obtained from regressions of DX. The only model in which a negative income effect occurs for DX is the original model of BF with the constrained sample (Table 21), again attesting to the BF sample constraints positively influencing the results.

Third, more sophisticated measures of commuting costs do not perform better than the two used by BF. Of all the regressions with the TAM measures across PDG and DX, only the change in the excess gallons of fuel consumed due to delay is significant in one model. Further, $Transit$ outperforms $Vehicles$ in every DX model, while neither seems to work consistently well with PDG. Lastly, neither removing the largest urban areas from the sample nor accounting for dual-income households seemed to consistently help explain sprawl. With regard to dual-income households, I may have chosen a poor proxy with female labor force participation; positive and significant coefficients are obtained only for the PDG in the year 2000. With regard to omitting large urban areas, if anything, the sample constraints originally imposed by BF produced stronger results, as expected, with DX, and qualitatively, though not quantitatively, stronger results with PDG. However, I question the empirical appropriateness of the BF sample constraints, as Nechyba and Walsh (2004) contend the Census defined urban area sufficiently delineates the built-up portion of the city.

6. Concluding Remarks

The literature has struggled for decades to accurately capture the process of sprawl in a scalar measure. One metric motivated by the Mills-Muth urban land use theory is the spatial size of an urban area. BF test the Mills-Muth model, and their results have been heavily relied upon as supporting the theory. In this article, I show that the results of BF cannot be replicated in- or out-of-sample and test two alternate measures of urban sprawl. The measure I develop, the directional index, outperforms the population density gradient, and thus provides urban economists with a new empirical tool that advances the study of urban sprawl. In that DX relaxes the strict monotonicity constrained implied by PDG and is measured *over time*, it is better able to capture the dynamic process of sprawl in a scalar measure. Additionally, I test the effects of dual-income households, removing the largest areas from the sample, and alternate measures of commuting costs, and I find no consistency across PDG and DX.

References

Alperovich, G. and Deutsche, J. (1994). Joint Estimation of Population Density Functions and the Location of the Central Business District. Journal of Urban Economics, 36, 239-248.

Anas, A. and Kim, I. (1996). General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration. Journal of Urban Economics, 40, 232-256.

Anas, A. and Xu, R. (1999). Congestion, Land Use, and Job Dispersion: A General Equilibrium Model. Journal of Urban Economics, 45, 451-473.

Box, G. and Cox, D. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society, Series B 26, 211-243.

Brueckner, J. and Fansler, D. (1983). The Economics of Urban Sprawl: Theory and Evidence on the Spatial Sizes of Cities. The Review of Economics and Statistics, 65, 479-482.

Brueckner, J. (1987). The structure of urban equilibria: a unified treatment of the Mills-Muth model. In E. Mills (Ed.), Handbook of Regional and Urban Economics (pp.703-1322). Amsterdam; New York: North-Holland.

Dudani, S. A. (1976). The Distance-Weighted k-Nearest-Neighbor Rule. IEEE Transaction on Systems, Man & Cybernetics, SMC6:4, 325-327.

Fujita, M., and Ogawa, H. (1982). Multiple Equilibria and Structural Transitions of Non-monocentric Urban Configurations. Regional Science and Urban Economics 12, 161-196.

Geoghegan, J., Wainger, L., and Bockstael, N. (1997). Spatial Landscape Indices in a Hedonic Framework: an Ecological Economics Analysis Using GIS. Ecological Economics 23:3, 251-264.

Glaeser, E. L. and Kahn, M. E. (2003). Sprawl and Urban Growth. NBER Working Paper No. w9733.

Glaeser, E.L. & J. Shapiro (2003). Urban Growth in the 1990s: Is City Living Back? Journal of Regional Science, 43, 139-165.

Henderson, V. and Mitra, A. (1996). The New Urban Landscape: Developers and Edge Cities. Regional Science and Urban Economics 26, 613-643.

Kau, J. B., and Lee, C. F. (1976). The Functional Form in Estimating the Density Gradient: An Empirical Investigation. Journal of the American Statistical Association 71, 326-327.

Lam, W. and Chong, T. T. (2006). Profitability of the Directional Indicators. Applied Financial Economics Letters, 2:6, 401-406.

LeRoy, S. F. and Sonstelie, J. (1983). Paradise Lost and Regained: Transportation Innovation, Income, and Residential Location. Journal of Urban Economics 13, 67-89.

Margo, R. (1992). Explaining the Postwar Suburbanization of Population in the United States: The Role of Income. Journal of Urban Economics, 31:2, 301-310.

Mills, E.S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. American Economic Review, 57:2, 197-210.

Mills, E.S, and Hamilton, B.W. (1994). Urban Economics, 5ed. New York: Harper-Collins.

Mills, E.S., and Lubuele, L.S. (1997). Inner Cities. Journal of Economic Literature, 35:2, 727-756.

Muth, R. F. (1969). Cities and Housing. Chicago: University of Chicago Press.

Nechyba, T. J. and Walsh, R. P. (2004). Urban Sprawl. Journal of Economic Perspectives, 18:4, 177-200

Redfearn, C. L. (2007). The Topography of Metropolitan Employment: Identifying Centers of Employment in a Polycentric Urban Area. Journal of Urban Economics, 61, 519-541.

Spivey, C. (2008). The Mills-Muth Model of Urban Spatial Structure: Surviving the Test of Time? Urban Studies 42: 2, 295-312.

Turnbull, G. K. (1998). Housing Prices and Residential Land Use under Job Site Uncertainty. Journal of Housing Economics, 7: 1, 1-20.

Turnbull, G. K. (n.d.). The Monocentric Urban Residential Land Market. Unpublished manuscript, University of Central Florida, Orlando, FL.

Wheaton, W.C. (1974). A Comparative Static Analysis of Urban Spatial Structure. Journal of Economic Theory, 9, 223-237.

Wilder, J. W. (1978). New concepts in technical trading systems (p. 70). McLeansville, NC: Trend Research.

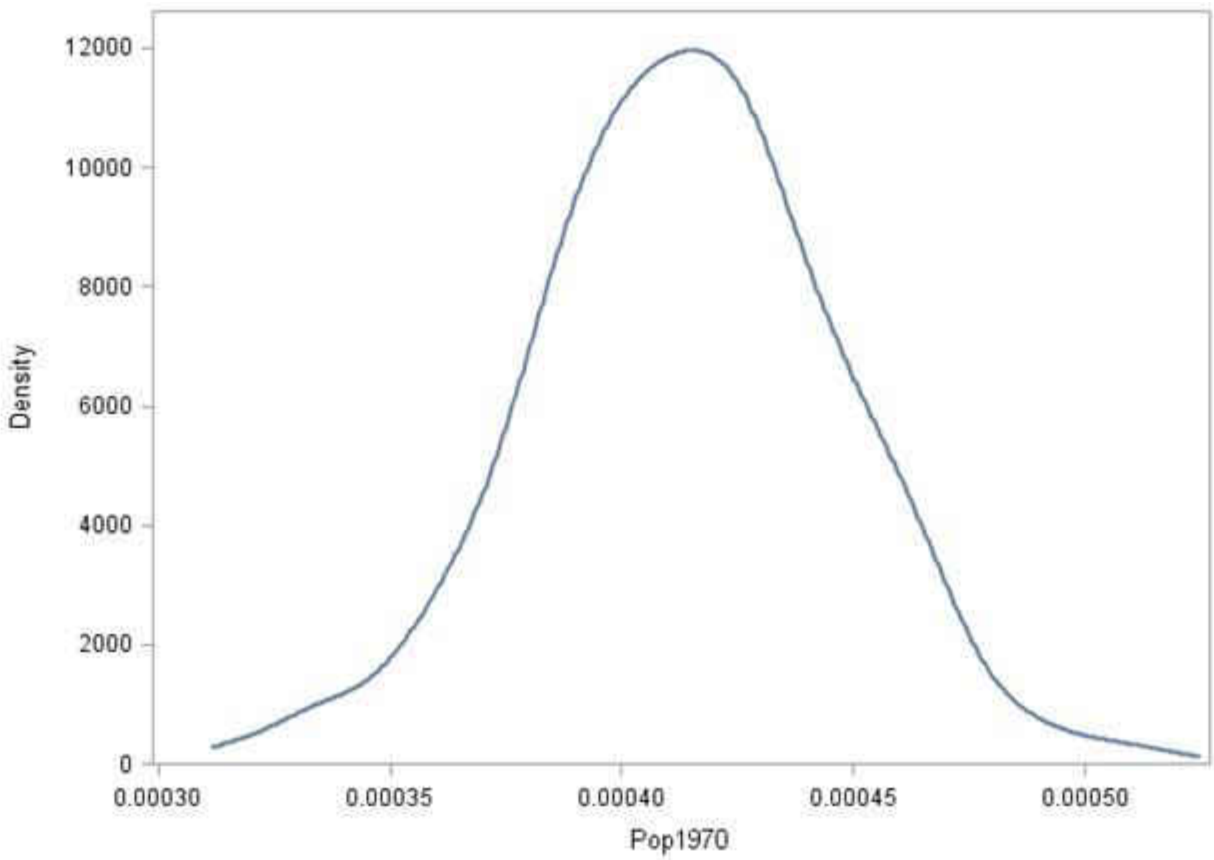

Figure 2: Kernel Density Estimate of 1970 Population (Model I)

Kernel density estimate of the OLS coefficient estimates of population from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Data were obtained from the 1970 U.S. Decennial Census. The data source was sampled 1,000 times, with each sample taken without replacement.

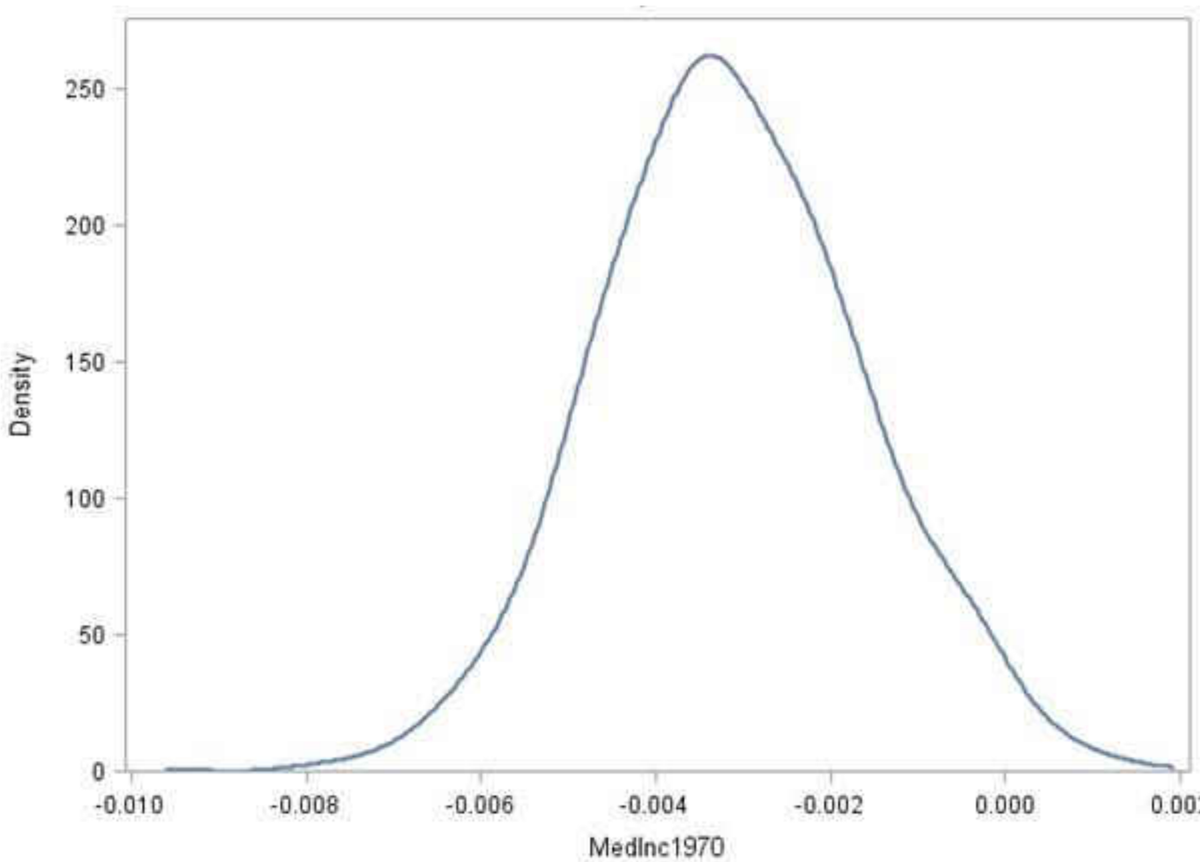

Figure 3: Kernel Density Estimate of 1970 Median Income (Model I)

Kernel density estimate of the OLS coefficient estimates of median income from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Data were obtained from the 1970 U.S. Decennial Census. The data source was sampled 1,000 times, with each sample taken without replacement.

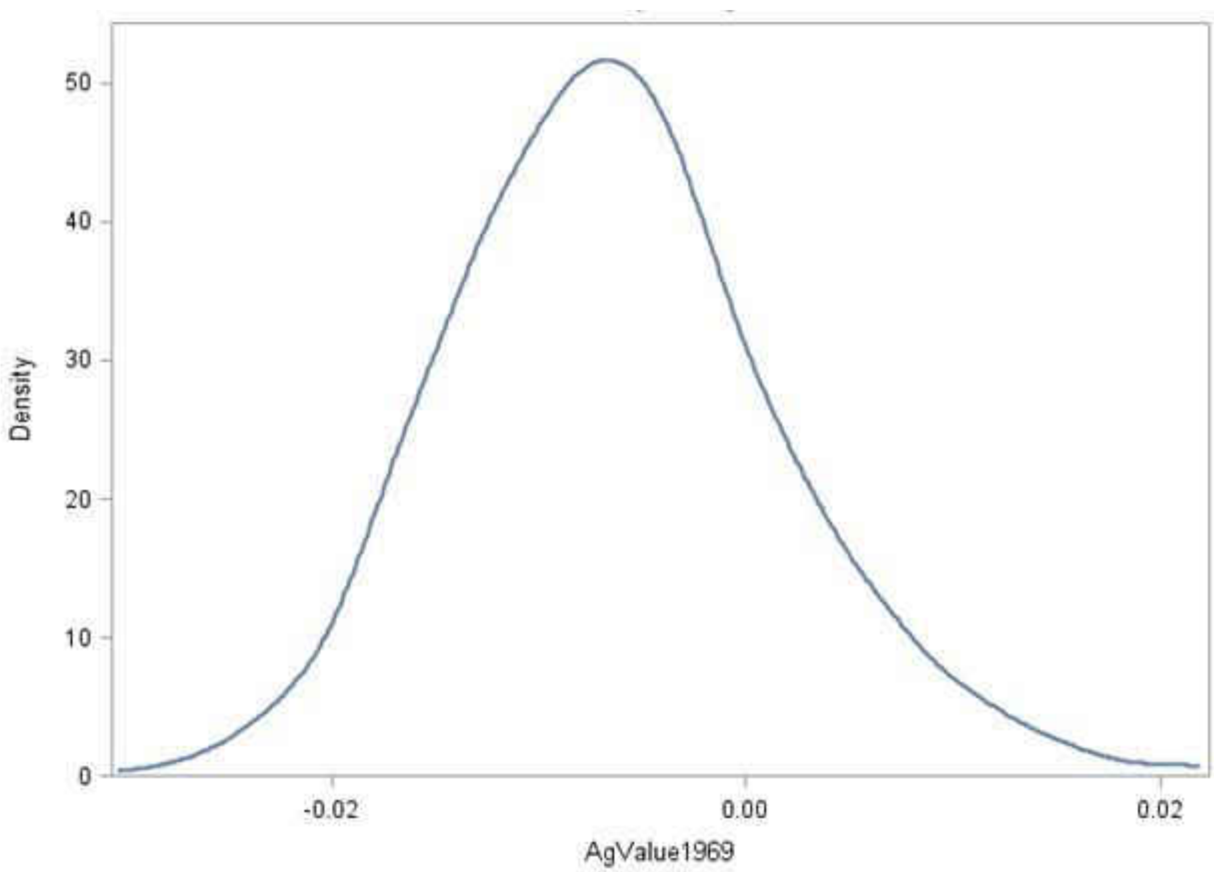

Figure 4: Kernel Density Estimate of 1969 Value of Agricultural Land (Model I)

Kernel density estimate of the OLS coefficient estimates of agricultural land value from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Agricultural land values were obtained from the 1969 USDA Census. The data source was sampled 1,000 times, with each sample taken without replacement.

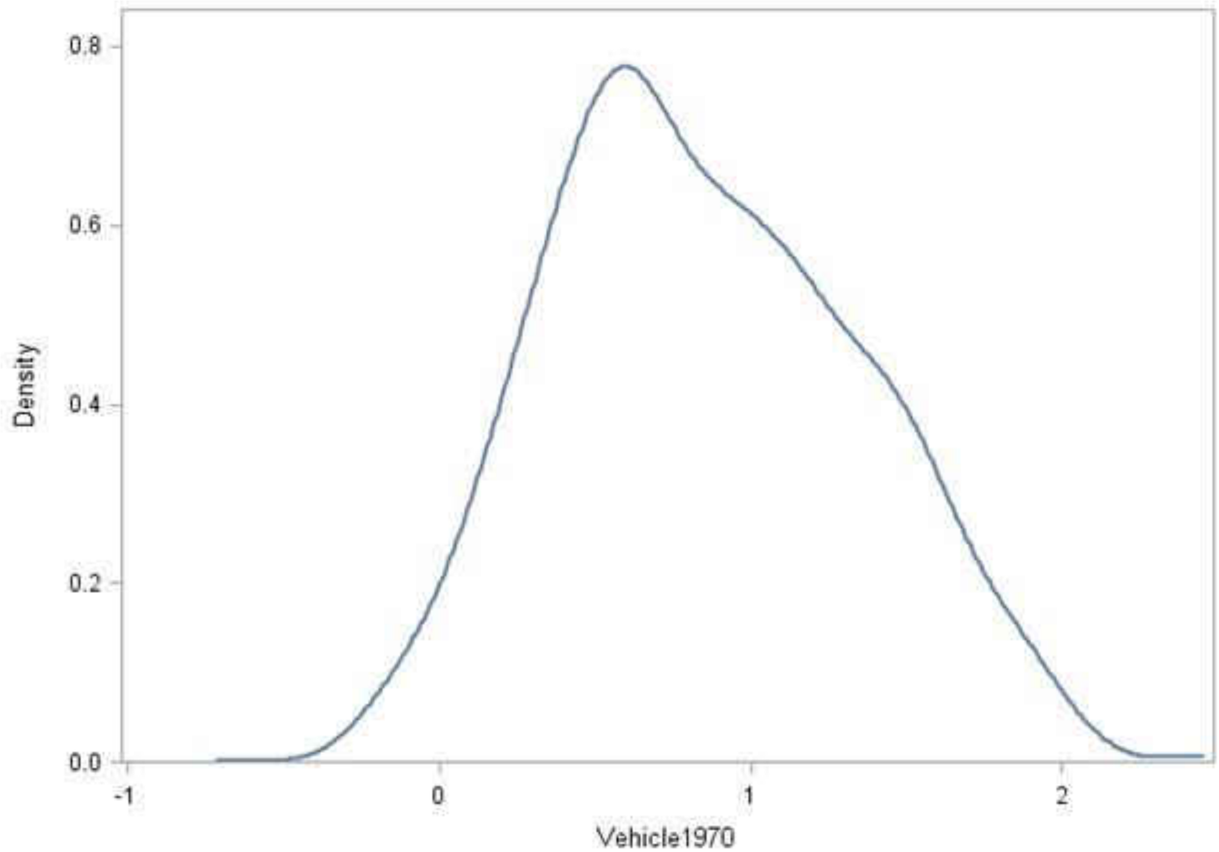

Figure 5: Kernel Density Estimate of 1970 Percent of Households with at least One Vehicle (Model I)

Kernel density estimate of the OLS coefficient estimates of the percent of households with at least one vehicle from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Data were obtained from the 1970 U.S. Decennial Census. The data source was sampled 1,000 times, with each sample taken without replacement.

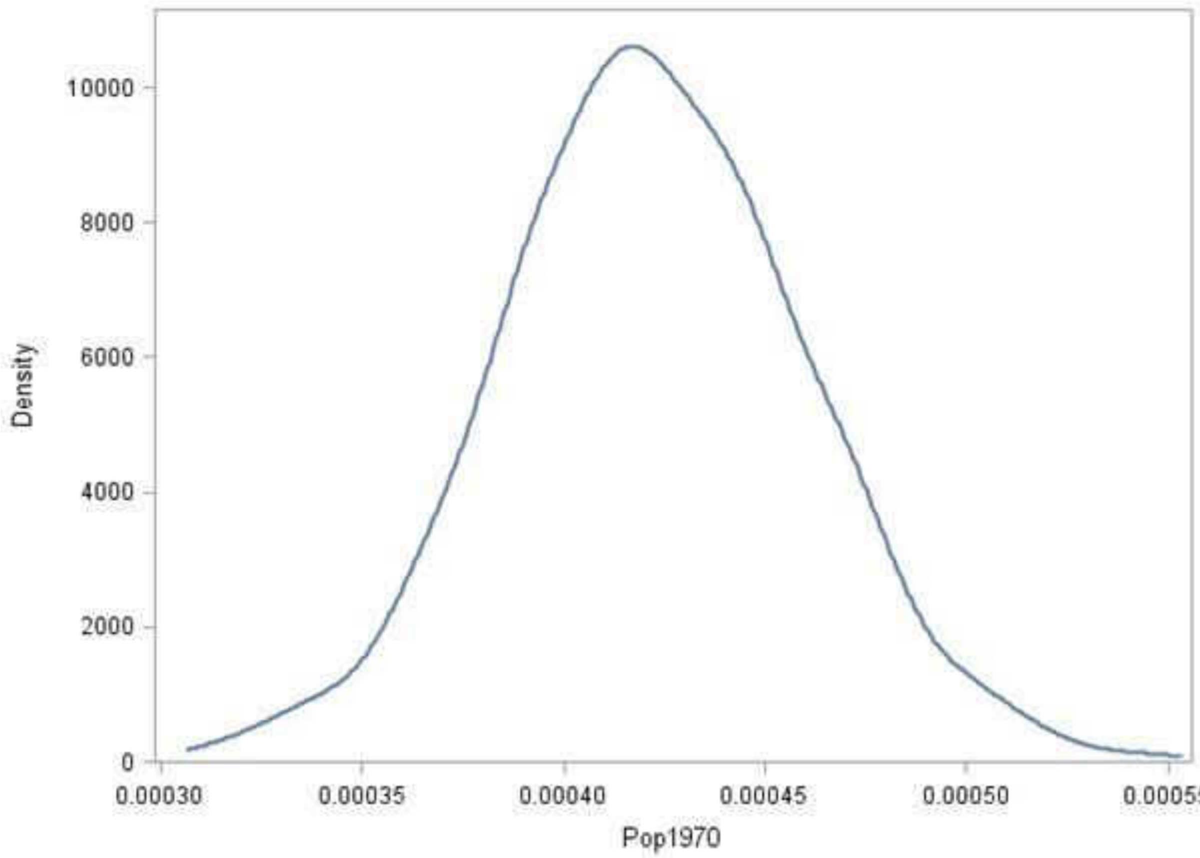

Figure 6: Kernel Density Estimate of 1970 Population (Model II)

Kernel density estimate of the OLS coefficient estimates of population from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Data were obtained from the 1970 U.S. Decennial Census. The data source was sampled 1,000 times, with each sample taken without replacement.

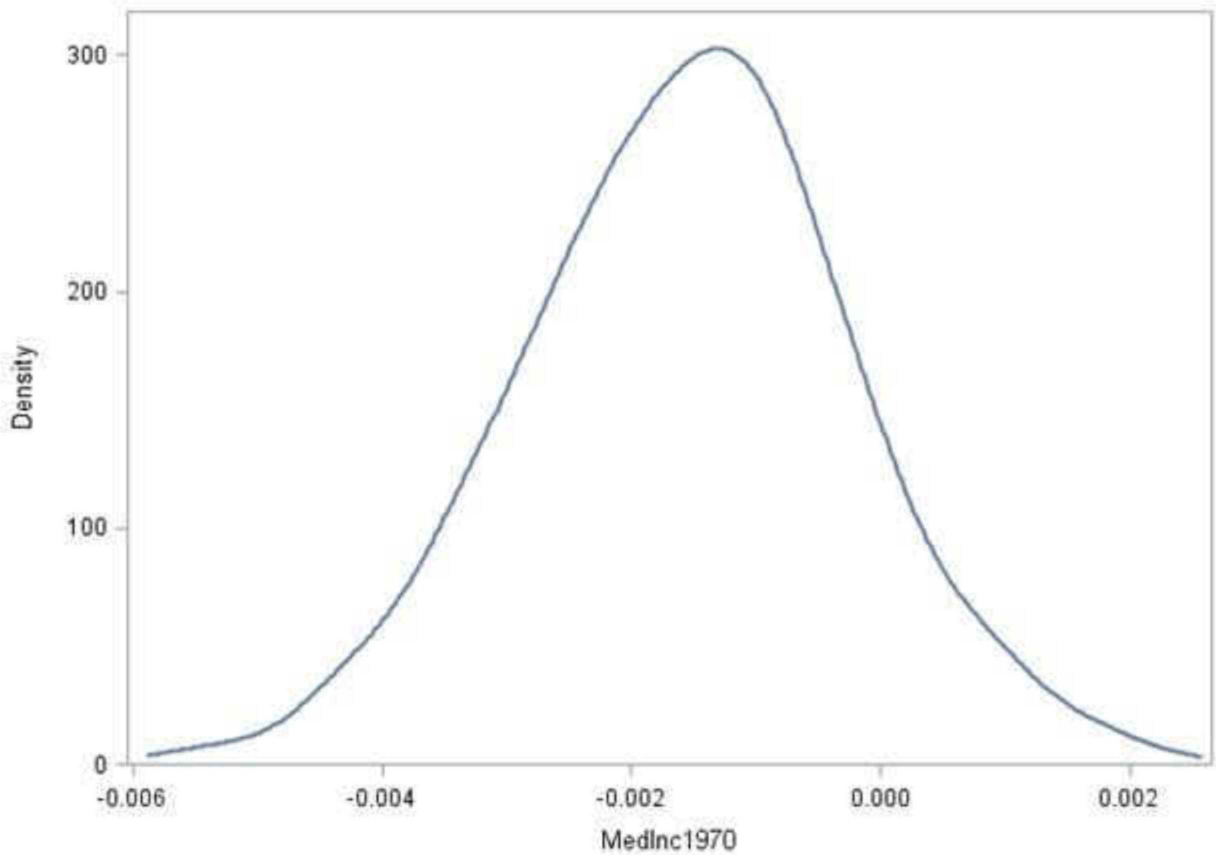

Figure 7: Kernel Density Estimate of 1970 Median Income (Model II)

Kernel density estimate of the OLS coefficient estimates of median income from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Data were obtained from the 1970 U.S. Decennial Census. The data source was sampled 1,000 times, with each sample taken without replacement.

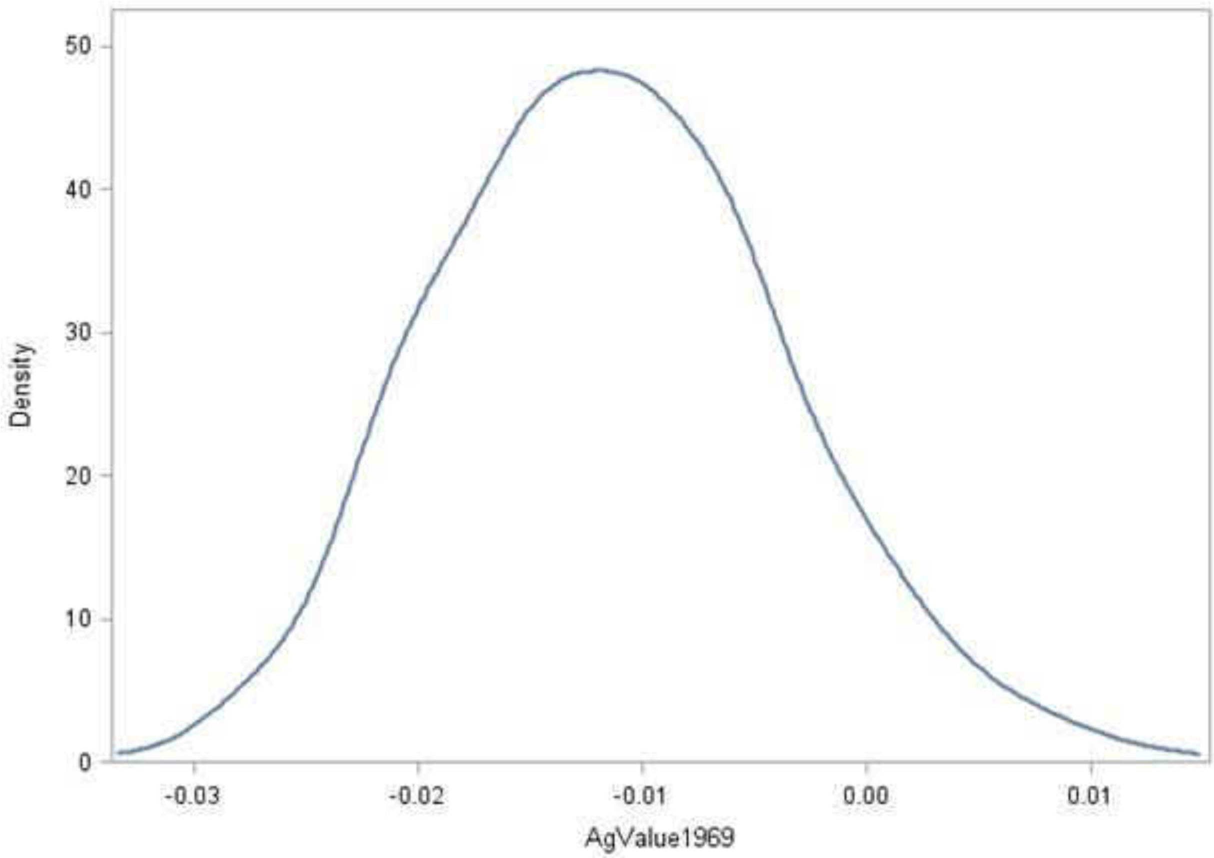

Figure 8: Kernel Density Estimate of 1969 Value of Agricultural Land (Model II)

Kernel density estimate of the OLS coefficient estimates of agricultural land value from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Agricultural land values were obtained from the 1969 USDA Census. The data source was sampled 1,000 times, with each sample taken without replacement.

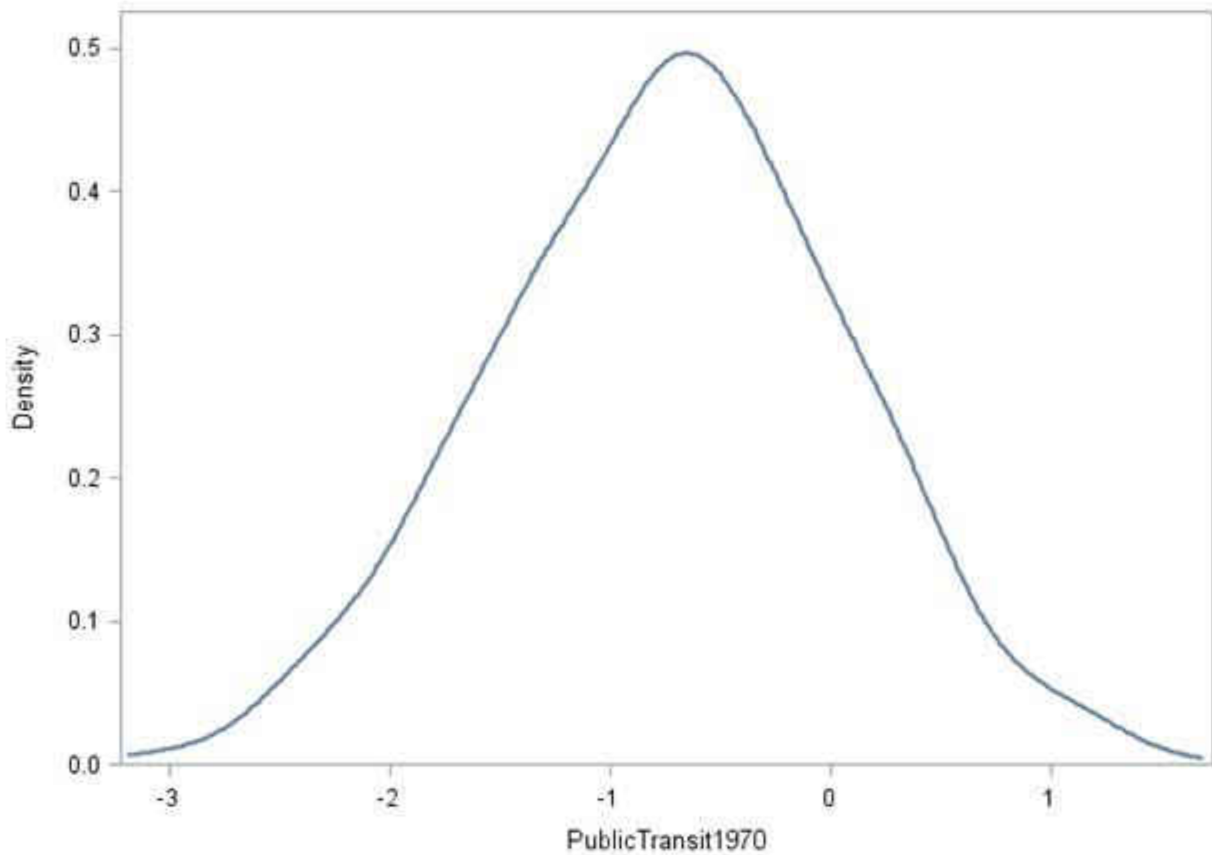

Figure 9: Kernel Density Estimate of 1970 Percent of Commuters using Public Transit, Excluding Taxis (Model II)

Kernel density estimate of the OLS coefficient estimates of the percent of commuters using public transit, excluding taxis, from repeated sampling of the 1970 data source used in Brueckner and Fansler (1983). Data were obtained from the 1970 U.S. Decennial Census. The data source was sampled 1,000 times, with each sample taken without replacement.

Table 12: Example Urban Areas

Population and land area changes for sample urban areas. Areas which experienced sprawl are presented in panel A, and areas which experienced contraction are presented in panel B. Land area is in square miles, and *DX* is the directional index, as computed by Equation 17. The deciles are calculated for each urban area based on the distance from the CBD centroid to the urban boundary.

Panel A: Urban Areas that Experienced Sprawl

Urban Area	Pocatello, ID			Tuscaloosa, AL		
	2000 Land Area: 22.70 2010 Land Area: 31.01 DX: -58.21			2000 Land Area: 91.95 2010 Land Area: 89.52 DX: -29.70		
Decile	2000 Population	2010 Population	Population Change	2000 Population	2010 Population	Population Change
1	5,952	5,803	-149	12,808	18,421	5,613
2	12,302	6,709	-5,593	12,449	11,801	-648
3	8,823	6,670	-2,153	17,435	16,428	-1,007
4	8,373	7,006	-1,367	19,913	14,045	-5,868
5	10,123	4,026	-6,097	30,512	15,735	-14,777
6	10,835	5,533	-5,302	19,281	16,150	-3,131
7	9,449	12,524	3,075	8,892	10,805	1,913
8	8,834	12,350	3,516	9,735	11,855	2,120
9	4,020	8,400	4,380	10,297	17,733	7,436
10	936	788	-148	2,504	6,141	3,637

Table 12: Example Urban Areas

Panel B: Urban Areas that Experienced Contraction

Urban Area		Danbury, CT			Missoula, MT	
		2000 Land Area: 59.33 2010 Land Area: 131.66 DX: 44.62			2000 Land Area: 35.75 2010 Land Area: 45.20 DX: 18.95	
Decile	2000 Population	2010 Population	Population Change	2000 Population	2010 Population	Population Change
1	10,643	18,058	7,415	6,974	6,026	-948
2	15,776	17,509	1,733	8,545	10,670	2,125
3	15,217	24,679	9,462	9,587	7,926	-1,661
4	21,928	25,460	3,532	8,262	7,757	-505
5	14,546	20,263	5,717	13,483	14,096	613
6	16,602	20,361	3,759	7,851	12,266	4,415
7	22,112	13,468	-8,644	5,553	10,834	5,281
8	15,966	5,495	-10,471	4,310	9,091	4,781
9	14,031	8,963	-5,068	11,455	141	-11,314
10	8,513	7,067	-1,446	9,227	3,350	-5,877

Table 13: Descriptive Statics for the Year 1970 Sample

Descriptive statistics for the year 1970 sample (N = 79) for the variables: land area of the urban area in square miles (*Land*), median income in dollars(*MedInc*), the percent of household with at least one vehicle (*Vehicles*), the percent of commuters that use public transit excluding taxis (*Transit*), the value of agricultural land per acre in dollars (*AgValue*), and urban area population (*Pop*). Data were obtained from the 1970 U.S. Decennial Census and the 1969 USDA Census.

Variable	Mean	Median	Minimum	Maximum	Standard Deviation
Land	46.29	38.70	11.90	121.20	24.99
MedInc	9,412	10	4,894	12,322	1,347
Vehicles	85.33	86.80	67.30	94.20	5.14
Transit	3.69	3.00	0.30	13.00	2.72
AgValue	406.00	370.00	33.00	959.00	218.00
Pop	118,660	103,300	52,627	255,824	54,672

Table 14: Descriptive Statics for the Year 2000 Population Density Gradient

Descriptive statics for the year 2000 unconstrained (N = 440) and constrained (N = 48) samples in panels A and B, respectively. The variables are: the population density gradient ($Gradient$), average rental value of agricultural land ($Avg. AgValue$), median income ($MedInc$), population (Pop), spatial area in square miles of the urban area ($Area$), percent of households with at least one vehicle ($Vehicles$), and percent of commuters using public transportation excluding taxis ($Transit$). Data were obtained from the 2000 U.S. Decennial Census.

Panel A: Unconstrained Sample

	Mean	Median	Minimum	Maximum	Standard Deviation
$Gradient$	0.06	0.46	-0.99	3.54	0.50
$Avg. AgValue$	4,095	3,029	149	29,392	3,748
$MedInc$	40,438	38,229	20,914	78,971	9,357
Pop	173,022	48,314	16,513	6,844,393	481,132
$Area$	160.82	63.43	12.10	3,352.60	305.67
$Vehicles$	91.26	91.70	67.60	97.23	3.05
$Transit$	1.80	1.10	0.00	28.20	2.47

Panel B: Constrained Sample

	Mean	Median	Minimum	Maximum	Standard Deviation
$Gradient$	0.40	0.36	0.02	1.26	0.25
$Avg. AgValue$	4,492	3,235	595	15,544	3,651
$MedInc$	41,882	40,359	24,964	71,537	10,027
Pop	95,679	83,557	52,668	209,030	39,292
$Area$	102.83	91.17	40.64	313.83	51.38
$Vehicles$	92.10	92.28	87.29	96.40	2.20
$Transit$	2.12	1.20	0.20	10.50	2.14

Table 15: Descriptive Statistics for the Year 2010 Population Density Gradient

Descriptive statics for the year 2000 unconstrained (N = 308) and constrained (N = 103) samples in panels A and B, respectively. The variables are: the population density gradient (*Gradient*), average rental value of agricultural land (*Avg. AgValue*), median income (*MedInc*), population (*Pop*), spatial area in square miles of the urban area (*Area*), percent of households with at least one vehicle (*Vehicles*), and percent of commuters using public transportation excluding taxis (*Transit*). Data were obtained from the 2010 U.S. Decennial Census.

Panel A: Unconstrained Sample

	Mean	Median	Minimum	Maximum	Standard Deviation
Gradient	0.47	0.35	-1.33	5.29	0.50
Avg. AgValue	6,422	5,151	209	58,587	5,644
MedInc	56,519	43,575	27,153	95,538	11,201
Pop	599,226	176,617	62,182	18,351,295	1,552,850
Area	228.77	91.68	14.12	3,450.20	396.35
Vehicles	92.37	92.55	68.62	97.63	2.71
Transit	2.01	1.20	0.00	31.30	2.71

Panel B: Constrained Sample

	Mean	Median	Minimum	Maximum	Standard Deviation
Gradient	0.78	0.66	-0.16	5.29	0.66
Avg. AgValue	5,597	4,682	209	26,446	4,543
MedInc	44,928	41,662	27,153	95,538	12,513
Pop	118,512	98,413	62,182	247,421	48,373
Area	57.14	49.91	14.17	182.28	28.66
Vehicles	92.74	92.77	86.56	96.74	2.18
Transit	1.70	1.10	0.20	8.90	1.71

Table 16: Descriptive Statistics for the Directional Index

Descriptive statistics for the unconstrained (N = 2,937) and constrained (N = 453) sample in panels A and B, respectively. The variables are: the directional index (DX), the year 2000 urban area population (Pop) and spatial area in square miles of the urban area ($Area$), and the changes in in median income ($\Delta MedInc$), the value of agricultural land ($\Delta AgValue$), the percent of households with at least one vehicle ($\Delta Vehicles$), and the percent of commuters using public transit excluding taxis ($\Delta Transit$). Data were obtained from the 2000 and 2010 U.S. Decennial Census programs.

Panel A: Unconstrained Sample

	Mean	Median	Minimum	Maximum	Standard Deviation
DX	23.24	83.61	-100.00	100.00	88.22
Pop	180,860	46,412	16,513	6,844,393	556,017
$Area$	176.06	54.14	12.10	3,352.60	342.33
$\Delta MedInc$	6,418	5,301	-43,667	59,827	14,580
$\Delta AgValue$	2,478	1,797	-55,302	199,032	10,082
$\Delta Vehicles$	1.13	0.91	-23.48	26.21	3.89
$\Delta Transit$	0.18	0.20	-28.10	29.80	3.43

Panel B: Constrained Sample

	Mean	Median	Minimum	Maximum	Standard Deviation
DX	-70.17	-100.00	-100.00	91.83	58.22
Pop	93,090	82,429	52,668	209,030	36,345
$Area$	102.32	90.40	40.64	313.83	52.01
$\Delta MedInc$	2,004	1,730	-33,696	41,220	12,713
$\Delta AgValue$	909	482	-12,177	23,504	6,449
$\Delta Vehicles$	0.44	0.16	-3.53	6.44	2.26
$\Delta Transit$	-0.07	-0.40	-8.30	8.00	2.95

Table 17: Hypothesis Tests

Hypothesis tests of the 1,000 1970 census repeated regression results for conformity with the Brueckner and Fansler (1983) empirical results. For each urban area, the variables are the population (Pop), median incomes ($MedInc$), rental value of agricultural land ($AgValue$), percent of households with at least one vehicle ($Vehicles$), and percent of commuters using public transportation, excluding taxis ($Transit$). Rejection of the null hypothesis at the 1% level is indicated with an *. Data were obtained from the 1970 U.S. Decennial Census and the 1969 USDA Census.

Variable	Null Hypothesis	t-statistic	Fail to Reject Null (%)	t-statistic	Fail to Reject Null (%)
Pop	$\mu \geq 0$	392.31	100.0	353.42	100.0
$MedInc$	$\mu \geq 0$	-64.95 *	2.0	-37.05 *	10.8
$AgValue$	$\mu \leq 0$	-25.34	80.7	-46.05	92.1
$Vehicles$	$\mu \geq 0$	53.86	97.1		
$Transit$	$\mu \leq 0$			-28.76	81.8

Table 18: Descriptive Statistics and Results from Repeated Sampling

Descriptive Statistics for the estimated coefficients from 1,000 repeated samples of the 1970 census data (panel A): median income in dollars (*MedInc*), the percent of households with at least one vehicle (*Vehicles*), the percent of commuters that use public transit excluding taxies (*Transit*), the value of agricultural land per acer in dollars (*AgValue)*, and urban area population (*Pop*). Reproduced results from Brueckner and Fansler (1983) linear regressions are reported (panel B). The reproduced results in panel B are expressed in scientific notation to remain consistent with results reported in panel A. Brueckner and Fansler (1983) reported their results to five decimal places. Data were obtained from the 2000 and 2010 U.S. Census programs, as discussed in Section 4. Data.

Dependent variable: square miles of land area

Panel A

Variable	Mean	Minimum	Maximum	Standard Deviation	Mean	Minimum	Maximum	Standard Deviation
*Pop**10^{-4}	4.14	3.11	5.25	0.33	4.22	3.06	5.54	0.38
*MedInc**10^{-3}	-3.15	-9.61	1.92	1.53	-1.56	-5.90	2.58	1.33
*AgValue**10^{-3}	-6.36	-30.30	2.19	7.94	-11.40	-33.40	14.80	7.81
Vehicles	0.85	-0.72	2.45	0.50				
Transit					-0.73	-3.19	1.69	0.80

Panel B	Coefficient Estimates	
*Pop**10^{-4}	4.00	4.10
*MedInc**10^{-3}	6.24	6.20
*AgValue**10^{-3}	- 28.88	-30.28
Vehicles	0.25	
Transit		-0.24

Table 19: Out-of-Sample Tests

OLS regression results from out-of-sample tests of the Brueckner and Fansler (1983) model with the unconstrained and constrained samples. For each urban area, the independent variables are median income (*Medinc*), population (*Pop*), the rental value of agricultural land per acer (*AgValue*), the percent of households with at least one vehicle (*Vehicles*), and the percent of commuters using public transit (*Transit*). Independent variables are for the appropriate year, either 2000 or 2010, except for the rental value of agricultural land, which is for the years 2002 or 2012, aligning with the years 2000 and 2010, respectively. The dependent variable in all regressions is the spatial size of the urban area in square miles, for the appropriate year, either 2000 or 2010. t-statistics are listed in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c, respectively.

Dependent variable: square miles of land area

	Unconstrained				Constrained			
Year	2000		2010		2000		2010	
Variable	N = 440		N = 309		N = 48		N = 103	
Intercept	-105.41	9.42	-509.80	85.60[b]	-289.59	59.83[a]	-92.56	20.84[b]
	(-0.56)	(0.39)	(-1.28)	(2.05)	(-1.31)	(2.80)	(-1.00)	(2.56)
*MedInc*10-4*	5.79	10.80[c]	-18.30	-5.87	-20.90[a]	-15.00[b]	-6.24[a]	-4.27[b]
	(0.77)	(1.66)	(-1.64)	(-0.59)	(-3.20)	(-2.60)	(-3.04)	(-2.57)
*Pop*10-4*	5.83[a]	5.94[a]	2.31[a]	2.36[a]	10.60[a]	10.70[a]	4.40[a]	4.48[a]
	(44.00)	(40.33)	(31.02)	(27.12)	(9.46)	(9.08)	(11.56)	(11.98)
*AgValue*10-3*	3.02[c]	3.50[c]	6.97[a]	7.15[a]	0.77	0.51	1.07[b]	1.25[a]
	(1.67)	(1.94)	(3.41)	(3.49)	(0.54)	(0.35)	(2.26)	(2.65)
Vehicles	1.42		6.89		4.07		1.29	
	(0.65)		(1.53)		(1.59)		(1.24)	
Transit		-5.07		-9.97[c]		0.27		-2.70[b]
		(-1.64)		(-1.82)		(0.12)		(-2.42)
Adj. R^2	87.2%	87.3%	82.4%	82.4%	66.9%	65.0%	59.9%	61.5%

Table 20: Population Density Gradient

OLS regression results of the population density gradient with the unconstrained and constrained samples. For each urban area, the independent variables are median income (*Medinc*), population (*Pop*), the rental value of agricultural land per acer (*AgValue*), the percent of households with at least one vehicle (*Vehicles*), and the percent of commuters using public transit (*Transit*). Independent variables are for the appropriate year, either 2000 or 2010, except for the rental value of agricultural land, which is for the years 2002 or 2012, aligning with the years 2000 and 2010, respectively. The dependent variable in all population density gradient for the appropriate year. t-statistics are listed in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c, respectively.

Dependent variable: population density gradient

	Unconstrained				Constrained			
Year	2000		2010		2000		2010	
Variable	N = 440		N = 309		N = 48		N = 103	
Intercept	0.71	0.91[a]	0.01	0.70[a]	2.15	0.97[a]	-1.07	1.32[a]
	(0.86)	(8.71)	(0.01)	(5.73)	(1.37)	(6.60)	(-0.33)	(4.51)
MedInc$*10^{-6}$	-7.21[b]	-8.04[a]	-3.95	-3.55	-5.08	-8.57[b]	-4.26	-1.36
	(-2.20)	(-2.85)	(-1.22)	(-1.23)	(-1.10)	(-2.16)	(-0.59)	(-0.23)
Pop$*10^{-7}$	-2.38[a]	-3.11[a]	-0.75[a]	-0.96[a]	-28.60[a]	-27.10[a]	-41.00[a]	-42.60[a]
	(-4.10)	(-4.84)	(-3.48)	(-3.78)	(-3.59)	(-3.32)	(-3.06)	(-3.16)
AgValue$*10^{-7}$	-9.29	-52.40	-32.20	52.20	18.10	20.60	31.60	12.00
	(-0.12)	(-0.66)	(-0.54)	(-0.88)	(0.18)	(-0.20)	(0.19)	(0.07)
Vehicles$*10^{-2}$	0.21		0.76		-1.39		2.70	
	(0.22)		(0.58)		(-0.76)		(0.74)	
Transit$*10^{-2}$		2.58[c]		1.31		1.61		0.73
		(1.92)		(0.83)		(1.00)		(0.18)
Adj. R^2	7.8%	8.5%	7.6%	7.7%	27.0%	27.8%	6.3%	5.8%

Table 21: Directional Index

OLS regression results of the directional index (*DX*) with the (un)constrained samples. The independent variables are the changes in median income (*Δincome*), the percent of households with at least one vehicle (*ΔVehicles*), the percent of commuters using public transportation, excluding taxis (*ΔTransit*), the rental value of agricultural land (*ΔAgValue*), and the percent of civilian labor force that is (*Female*). Either the year 2000 urban area population (*Pop*) or spatial size (*Area*) in square miles is used as a control variable. *t*-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% levels is denoted by a, b or c, respectively.

Dependent Variable: Directional Index

	Unconstrained				Constrained			
Variable	N = 2,937				N = 453			
Intercept	22.08[a]	30.61[a]	21.81[a]	30.24[a]	-15.85[b]	-42.21[a]	-12.97[b]	-39.93[a]
	(12.66)	(16.88)	(12.53)	(16.62)	(-2.57)	(-7.95)	(-2.21)	(-7.97)
*Δincome*10⁻⁴*	12.50[a]	10.70[a]	9.55[a]	9.49[a]	-18.50[a]	-15.80[a]	-18.70[a]	-17.10[a]
	(10.15)	(8.85)	(8.03)	(8.26)	(-7.11)	(-5.74)	(8.86)	(7.61)
ΔVehicles	-0.97[b]	-0.26			0.28	-0.62		
	(-2.10)	(-0.59)			(0.23)	(-0.47)		
ΔTransit			2.52[a]	1.24[b]			5.37[a]	5.35[a]
			(4.20)	(2.19)			(6.61)	(6.18)
*ΔAgValue*10⁻⁴*	7.11[a]	5.74[a]	6.16[a]	5.13[a]	54.20[a]	50.80[a]	43.40[a]	40.30[a]
	(4.28)	(3.55)	(3.69)	(3.15)	(12.39)	(11.00)	(9.72)	(8.49)
*Pop*10⁻⁵*	-4.17[a]		-3.69[a]		-59.80[a]		-61.26[a]	
	(-13.74)		(-11.30)		(-9.74)		(-10.45)	
Area		-0.09[a]		-0.09[a]		-0.28[a]		-0.29[a]
		(-19.22)		(-17.18)		(-6.21)		(-6.77)
Adj. R^2	15.6%	20.3%	16.1%	20.4%	34.0%	26.3%	39.3%	32.1%

Table 22: Population Density Gradient with Alternative Commuting Cost Measures (Year 2000)

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are median income (*MedInc*), population (*Pop*), rental value of agricultural land per acre (*AgValue*), freeway miles driven per auto commuter (*Freeway*), arterial street miles driven per auto commuter (*Arterial*), annual excess fuel consumed due to traffic congestion per auto commuter (*Gallons*), vehicle hours of delay due to traffic congestion per auto commuter (*Delay*), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (*TTI*), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period (*Stress*), and annual vehicle and time congestion cost in dollars (*Congestion*). Agricultural land rent values are for the year 2002, aligning with the year 2000. *t*-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c superscripts, respectively.

Dependent Variable: Population Density Gradient (2000)

Variable	N = 97						
Intercept	0.54^{a} (5.32)	0.52^{a} (4.04)	0.55^{a} (5.32)	0.54^{a} (5.38)	0.94^{a} (2.93)	0.60^{a} (2.68)	0.55^{a} (5.46)
*MedInc*10⁻⁶*	-5.99^{b} (-2.27)	-6.92^{a} (-2.78)	-6.19^{b} (-2.38)	-4.72^{c} (-1.71)	-5.29 (-1.33)	-6.54^{b} (-2.39)	-4.88^{c} (-1.83)
*Pop*10⁻⁸*	-5.10^{b} (-2.30)	-5.16^{b} (-2.30)	-4.00 (-1.58)	-2.52 (-0.95)	-3.57 (-1.44)	-5.85^{b} (-2.06)	-2.04 (-0.76)
*AgValue*10⁻⁷*	-7.00 (-0.18)	-7.47 (-0.18)	-8.29 (-0.21)	-15.30 (-0.39)	-8.57 (-0.22)	-7.78 (-0.19)	-10.80 (-0.28)
*Freeway*10⁻³*	-2.67 (-1.03)						
*Arterial*10⁻⁴*		6.33 (0.17)					
*Gallons*10⁻³*			-3.84 (0.91)				
*Delay*10⁻³*				-3.25^{c} (-1.75)			
TTI					-0.41 (-1.33)		
Stress						-0.07 (-0.34)	
*Congestion*10⁻⁴*							-1.34^{c} (-1.94)
Adj. R^2	18.3%	17.3%	18.1%	20.00%	18.9%	17.4%	20.6%

Table 23: Population Density Gradient with Alternative Commuting Cost Measures (Year 2010)

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are median income (*MedInc*), population (*Pop*), rental value of agricultural land per acre (*AgValue*), freeway miles driven per auto commuter (*Freeway*), arterial street miles driven per auto commuter (*Arterial*), annual excess fuel consumed due to traffic congestion per auto commuter (*Gallons*), vehicle hours of delay due to traffic congestion per auto commuter (*Delay*), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (TTI), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period (*Stress*), and annual vehicle and time congestion cost in dollars (*Congestion*). Agricultural land rent values are for the year 2012, aligning with the year 2010. *t*-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c superscripts, respectively.

Dependent Variable: Population Density Gradient (2010)

Variable	N = 83						
Intercept	0.33[a]	0.28[c]	0.31[b]	0.33[a]	0.37	0.23	0.34[a]
	(2.81)	(1.81)	(2.49)	(2.69)	(0.91)	(0.83)	(2.76)
$MedInc*10^{-6}$	-1.32	-1.73	-2.09	-1.89	-1.76	-2.32	-1.54
	(-0.46)	(-0.65)	(-0.76)	(-0.67)	(-0.60)	(-0.80)	(-0.55)
$Pop*10^{-8}$	-3.46[a]	-3.50[a]	-3.65[a]	-3.50[a]	-3.44[a]	-3.61[a]	-3.24[a]
	(-4.38)	(-4.44)	(-4.00)	(-3.55)	(-3.78)	(-4.24)	(-3.20)
$AgValue*10^{-7}$	-8.31	-9.37	-9.61	-8.60	-8.33	-10.20	-7.80
	(-0.29)	(-0.32)	(-0.33)	(-0.30)	(-0.29)	(-0.35)	(-0.27)
$Freeway*10^{-3}$	-1.85						
	(-0.52)						
$Arterial*10^{-3}$		2.41					
		(0.52)					
$Gallons*10^{-3}$			1.71				
			(0.37)				
$Delay*10^{-5}$				5.89			
				(0.03)			
TTI					-0.04		
					(-0.09)		
Stress						0.10	
						(0.38)	
$Congestion*10^{-5}$							-3.69
							(-0.38)
Adj. R^2	22.9%	22.9%	22.7%	22.6%	22.6%	22.8%	22.7%

Table 24: Directional Index with Alternative Commuting Cost Measures

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are changes in median income ($\Delta MedInc$), rental value of agricultural land per acre ($\Delta AgValue$), freeway miles driven per auto commuter ($\Delta Freeway$), arterial street miles driven per auto commuter ($\Delta Arterial$), annual excess fuel consumed due to traffic congestion per auto commuter ($\Delta Gallons$), vehicle hours of delay due to traffic congestion per auto commuter ($\Delta Delay$), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (ΔTTI), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period ($\Delta Stress$), and annual vehicle and time congestion cost in dollars ($\Delta Congestion$). Population (Pop) is the 2000 population level. t-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% level sis indicated by a, b, or c superscripts, respectively.

Dependent Variable: Directional Index

Variable	N = 803						
Intercept	74.82[a]	63.31[a]	81.29[a]	79.61[a]	79.87[a]	80.19[a]	74.84[a]
	(30.87)	(22.44)	(24.99)	(29.75)	(29.43)	(30.09)	(29.31)
$\Delta income*10^{-4}$	12.80[a]	13.80[a]	13.00[a]	13.00[a]	13.20[a]	12.70[a]	13.60[a]
	(7.92)	(8.60)	(7.88)	(7.96)	(8.12)	(7.77)	(8.20)
$\Delta AgValue*10^{-5}$	2.48	5.00	-4.97	-5.19	-5.47	-5.06	-5.74
	(0.22)	(-0.45)	(-0.44)	(-0.46)	(-0.48)	(-0.45)	(-0.50)
$Pop*10^{-5}$	-6.94[a]	-6.82[a]	-6.50[a]	-6.42[a]	-6.50[a]	-6.34[a]	-6.68[a]
	(-19.48)	(-19.27)	(-17.67)	(-17.68)	(-17.97)	(-17.49)	(-18.29)
$\Delta Freeway$	-1.01[a]						
	(-6.57)						
$\Delta Arterial$		-1.08[a]					
		(-7.01)					
$\Delta Gallons$			-0.18[a]				
			(-3.60)				
$\Delta Delay$				-0.42[a]			
				(-5.42)			
ΔTTI					-0.45[a]		
					(-5.25)		
$\Delta Stress$						-0.58[a]	
						(-5.98)	
$\Delta Congestion$							-0.02[c]
							(-1.94)
Adj. R^2	39.7%	40.1%	37.5%	38.7%	38.6%	39.2%	36.8%

Table 25: Directional Index with Alternative Commuting Cost Measures

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are changes in median income (*ΔMedInc*), rental value of agricultural land per acre (*ΔAgValue*), freeway miles driven per auto commuter (*ΔFreeway*), arterial street miles driven per auto commuter (*ΔArterial*), annual excess fuel consumed due to traffic congestion per auto commuter (*ΔGallons*), vehicle hours of delay due to traffic congestion per auto commuter (*ΔDelay*), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (*ΔTTI*), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period (*ΔStress*), and annual vehicle and time congestion cost in dollars (*ΔCongestion*). Area (*Area*) is the 2000 urban area spatial size in square miles. t-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% level sis indicated by a, b, or c superscripts, respectively.

Dependent Variable: Directional Index

Variable	N = 803						
Intercept	85.38^{a}	71.81^{a}	80.01^{a}	83.14^{a}	83.56^{a}	83.39^{a}	83.79^{a}
	(34.76)	(25.79)	(25.46)	(31.20)	(30.94)	(31.35)	(32.22)
$\Delta income * 10^{-4}$	9.98^{a}	11.30^{a}	11.60^{a}	11.40^{a}	11.40^{a}	11.30^{a}	11.20^{a}
	(6.44)	(7.33)	(7.20)	(7.06)	(7.06)	(7.03)	(6.99)
$\Delta AgValue * 10^{-6}$	56.49	8.12	3.56	-5.18	-7.06	-5.92	-25.20
	(0.53)	(0.76)	(0.03)	(-0.05)	(-0.06)	(-0.05)	(-0.23)
Area	-0.13^{a}	-0.12^{a}	-0.12^{a}	-0.12^{a}	-0.12^{a}	-0.12^{a}	-0.12^{a}
	(-22.66)	(-22.24)	(-19.59)	(-18.90)	(-19.25)	(-18.48)	(-20.50)
ΔFreeway	-1.23^{a}						
	(-8.42)						
ΔArterial		-1.24^{a}					
		(-8.41)					
ΔGallons			0.09^{a}				
			(1.65)				
$\Delta delay * 10^{-3}$				-8.42			
				(-0.10)			
ΔTTI					-0.05		
					(-0.56)		
ΔStress						-0.04	
						(-0.42)	
ΔCongestion							-0.02
							(-1.35)
Adj. R^2	45.9%	45.9%	41.3%	41.1%	41.1%	41.1%	41.2%

Table 26: Population Density Gradient with Workforce Demographics

OLS regression results of the population density gradient with the unconstrained and constrained samples. For each urban area, the independent variables are median income (*Medinc*), population (*Pop*), the rental value of agricultural land per acer (*AgValue*), the percent of households with at least one vehicle (*Vehicles*), the percent of commuters using public transit (*Transit*), and the percent of civilian labor force that is (*Female*). Independent variables are for the appropriate year, either 2000 or 2010, except for the rental value of agricultural land, which is for the years 2002 or 2012, aligning with the years 2000 and 2010, respectively. The dependent variable in all population density gradient for the appropriate year. *t*-statistics are listed in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c, respectively.

Dependent Variable: Population Density Gradient

	Unconstrained				Constrained			
Year	2000		2010		2000		2010	
Variable	N = 440		N = 308		N = 48		N = 102	
Intercept	0.24	0.40[c]	-0.25	0.43	0.79	0.14	-1.07	1.37[b]
	(0.29)	(1.70)	(-0.21)	(1.90)	(0.55)	(0.54)	(-0.31)	(2.32)
*MedInc*10⁻⁶*	-11.05[a]	-11.60[a]	-4.06	-3.64	-12.39[a]	-14.25[a]	-4.28	-1.51
	(-3.11)	(-3.68)	(-1.25)	(-1.26)	(-2.71)	(-3.69)	(-0.58)	(-0.24)
*Pop*10⁻⁷*	-2.41[a]	-3.02[a]	-0.76[a]	-0.96[a]	-22.00[a]	-21.30[a]	-41.00[a]	-42.50[a]
	(-4.18)	(-4.72)	(-3.52)	(-3.79)	(-3.02)	(-2.87)	(-3.01)	(-3.09)
*AgValue*10⁻⁶*	2.26	-1.51	-3.00	-4.95	12.49	12.53	3.16	1.36
	(0.29)	(-0.19)	(-0.51)	(-0.83)	(1.31)	(1.32)	(0.19)	(0.08)
*Vehicles*10⁻³*	1.08		7.49		-7.72		27.01	
	(0.11)		(0.57)		(-0.48)		(0.72)	
*Transit*10⁻³*		22.00		12.58		8.79		8.02
		(1.63)		(0.79)		(0.61)		(0.19)
*Female*10⁻³*	12.03[a]	11.20[b]	5.00	4.84	17.00[a]	16.82[a]	0.02	-0.75
	(2.67)	(2.48)	(-1.45)	(1.40)	(3.60)	(3.55)	(0.00)	(-0.08)
Adj. R^2	9.0%	9.6%	7.9%	8.0%	42.9%	43.1%	5.3%	4.8%

Table 27: Directional Index with Workforce Demographics

OLS regression results of the directional index (*DX*) with the (un)constrained samples. The independent variables are the changes in median income (*Δincome*), the percent of households with at least one vehicle (*ΔVehicles*), the percent of commuters using public transportation, excluding taxis (*ΔTransit*), the rental value of agricultural land (*ΔAgValue*), and the percent of civilian labor force that is (*ΔFemale*). Either the year 2000 urban area population (*Pop*) or spatial size in square miles (*Area*) is used as a control variable. *t*-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% levels is denoted by a, b or c, respectively.

Dependent Variable: Directional Index

	Unconstrained				Constrained			
Variable	N = 2,917				N = 453			
Intercept	22.78[a]	31.36[a]	22.61[a]	31.04[a]	12.13[a]	11.64	6.94	9.27
	(12.42)	(16.53)	(12.36)	(16.33)	(1.72)	(1.59)	(0.96)	(1.30)
*Δincome*10⁻⁴*	12.70[a]	10.90[a]	9.92[a]	9.83[a]	4.28	5.28[c]	1.03	-6.52
	(10.06)	(8.79)	(8.16)	(8.39)	(1.35)	(1.66)	(0.36)	(-0.23)
ΔVehicles	-0.90[c]	-0.19			-0.85	-1.50		
	(-1.93)	(-0.43)			(-0.71)	(-1.23)		
ΔTransit			2.49[a]	1.24[a]			5.69[a]	10.02[a]
			(4.16)	(2.17)			(2.75)	(5.09)
*ΔAgValue*10⁻⁴*	7.12[a]	5.84[a]	6.23[a]	5.20[a]	11.50	2.85	12.60[c]	5.38
	(4.33)	(3.61)	(3.73)	(3.19)	(1.55)	(-0.38)	(1.75)	-0.75
*Pop*10⁻⁵*	-4.21[a]		-3.73[a]		-55.84[a]		-51.45[a]	
	(-13.85)		(-11.40)		(-10.38)		(-9.20)	
Area		-0.09[a]		-0.09[a]		-0.45[a]		-0.43[a]
		(-19.31)		(-17.24)		(-9.85)		(-9.68)
ΔFemale	-0.19	-0.19	-0.21	-0.20	-0.16	-0.47[c]	-0.25	-0.61[b]
	(-1.20)	(-1.24)	(-1.34)	(-1.35)	(-0.58)	(-1.74)	(-0.92)	(-2.30)
Adj. R^2	15.7%	20.4%	16.1%	20.5%	17.3%	16.00%	18.3%	19.5%

Table 28: Population Density Gradient with Alternative Commuting Cost Measures and Workforce Demographics (Year 2000)

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are median income (*MedInc*), population (*Pop*), rental value of agricultural land per acre (*AgValue*), the percent of civilian labor force that is (*Female*), freeway miles driven per auto commuter (*Freeway*), arterial street miles driven per auto commuter (*Arterial*), annual excess fuel consumed due to traffic congestion per auto commuter (*Gallons*), vehicle hours of delay due to traffic congestion per auto commuter (*Delay*), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (*TTI*), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period (*Stress*), and annual vehicle and time congestion cost in dollars (*Congestion*). Agricultural land rent values are for the years 2002, aligning with the years 2000. *t*-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% level sis indicated by a, b, or c superscripts, respectively.

Dependent Variable: Population Density Gradient (2000)

Variable	N = 97						
Intercept	0.19	0.26	0.22	0.21	0.65^{c}	0.25	0.21
	(1.03)	(1.30)	(1.20)	(1.13)	(1.79)	(0.82)	(1.12)
$MedInc*10^{-6}$	-9.55^{a}	-10.20^{a}	-9.66^{a}	-8.17^{b}	-8.50^{b}	-10.20^{a}	-8.50^{a}
	(-3.12)	(-3.28)	(-3.14)	(-2.60)	(-2.55)	(-2.95)	(-2.77)
$Pop*10^{-8}$	-4.48^{b}	-4.65^{b}	-2.81	-1.36	-3.22	-4.67^{b}	-0.75
	(-2.04)	(-2.08)	(-1.10)	(-0.51)	(-1.30)	(-2.00)	(-0.28)
$AgValue*10^{-7}$	9.95	3.74	6.76	-1.69	3.22	3.45	4.33
	(0.25)	(0.09)	(0.17)	(-0.04)	(0.08)	(0.09)	(0.11)
$Female*10^{-3}$	8.81^{b}	6.84^{c}	8.24^{b}	8.39^{b}	6.59^{c}	6.90^{c}	8.67^{b}
	(2.17)	(1.73)	(2.04)	(2.14)	(1.68)	(1.70)	(2.22)
$Freeway*10^{-3}$	-4.36						
	(-1.65)						
$Arterial*10^{-4}$		1.21					
		(0.03)					
$Gallons*10^{-3}$			-5.99				
			(-1.40)				
$Delay*10^{-3}$				-3.97^{b}			
				(-2.15)			
TTI					-0.39		
					(-1.25)		
$Stress*10^{-3}$						9.72	
						(0.05)	
$Congestion*10^{-4}$							-1.64^{b}
							(-2.38)
Adj. R^2	21.4%	19.1%	20.8%	23.0%	20.5%	19.1%	23.8%

Table 29: Population Density Gradient with Alternative Commuting Cost Measures and Workforce Demographics (Year 2010)

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are median income (*MedInc*), population (*Pop*), rental value of agricultural land per acre (*AgValue*), freeway miles driven per auto commuter (*Freeway*), arterial street miles driven per auto commuter (Arterial), annual excess fuel consumed due to traffic congestion per auto commuter (*Gallons*), vehicle hours of delay due to traffic congestion per auto commuter (*Delay*), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (*TTI*), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period (*Stress*), and annual vehicle and time congestion cost in dollars (*Congestion*). Agricultural land rent values are for the year 2012, aligning with the year 2010. *t*-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c superscripts, respectively.

Dependent Variable: Population Density Gradient (2010)

Variable	N = 82						
Intercept	0.32^{b}	0.27	0.32^{b}	0.32^{b}	0.43	0.25	0.32^{b}
	(1.97)	(1.52)	(1.99)	(2.00)	(1.26)	(1.01)	(2.00)
MedInc * 10^{-6}	-4.00	-3.63	-3.87	-3.81	-3.40	-4.13	-3.73
	(-1.60)	(-1.57)	(-1.65)	(-1.56)	(-1.36)	(-1.65)	(-1.54)
Pop * 10^{-8}	-2.98^{a}	-3.00^{a}	-3.10^{a}	-3.02^{a}	-2.80^{a}	-3.12^{a}	-2.96^{a}
	(-4.09)	(-4.12)	(-3.54)	(-3.25)	(-3.19)	(-3.81)	(-3.07)
AgValue * 10^{-8}	13.74	6.30	8.63	12.68	19.40	1.70	14.44
	(0.06)	(0.03)	(0.03)	(0.05)	(0.08)	(0.01)	(0.06)
Female * 10^{-3}	1.38	1.33	1.20	1.33	1.62	1.07	1.38
	(0.59)	(0.56)	(0.49)	(0.55)	(0.66)	(0.43)	-0.56
Freeway * 10^{-4}	7.64						
	(0.25)						
Arterial * 10^{-3}		2.25					
		(0.57)					
Gallons * 10^{-3}			1.03				
			(0.25)				
Delay * 10^{-4}				1.31			
				(0.06)			
TTI					-0.13		
					(0.38)		
Stress						0.09	
						(0.38)	
Congestion * 10^{-6}							-2.83
							(-0.03)
Adj. R^2	27.7%	28.0%	27.7%	27.7%	27.8%	27.8%	27.7%

Table 30: Directional Index with Alternative Commuting Cost Measures and Workforce Demographics

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are changes in median income ($\Delta MedInc$), rental value of agricultural land per acre ($\Delta AgValue$), the percent of civilian labor force that is ($\Delta Female$), freeway miles driven per auto commuter ($\Delta Freeway$), arterial street miles driven per auto commuter ($\Delta Arterial$), annual excess fuel consumed due to traffic congestion per auto commuter ($\Delta Gallons$), vehicle hours of delay due to traffic congestion per auto commuter ($\Delta Delay$), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (ΔTTI), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period ($\Delta Stress$), and annual vehicle and time congestion cost in dollars ($\Delta Congestion$). Population (Pop) is the 2000 population level. t-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% level is indicated by a, b, or c superscripts, respectively.

Dependent Variable: Directional Index

Variable	N = 803						
Intercept	77.59[a]	64.91[a]	84.28[a]	83.14[a]	83.47[a]	82.99[a]	76.78[a]
	(30.39)	(22.70)	(24.28)	(28.85)	(28.50)	(29.28)	(28.27)
*Δincome * 10^{-3}*	1.40[a]	1.50[a]	1.39[a]	1.41[a]	1.44[a]	1.37[a]	1.43[a]
	(8.48)	(9.10)	(8.23)	(8.50)	(8.66)	(8.22)	(8.47)
*ΔAgValue * 10^{-5}*	6.57	8.67	-2.66	-2.12	2.47	-2.28	-4.12
	(0.59)	(-0.77)	(-0.23)	(-0.19)	(-0.22)	(-0.20)	(-0.36)
*Pop * 10^{-5}*	-7.16[a]	-7.00[a]	-6.61[a]	-6.56[a]	-6.64[a]	-6.47[a]	-6.79[a]
	(-19.87)	(-19.58)	(-17.89)	(-18.04)	(-18.33)	(-17.78)	(-18.44)
ΔFemale	-0.64[a]	-0.57[a]	-0.48[a]	-0.63[a]	-0.63[a]	-0.54[a]	-0.41[b]
	(-3.28)	(-2.96)	(-2.42)	(-3.18)	(-3.17)	(-2.81)	(-2.07)
ΔFreeway	-1.12[a]						
	(-7.18)						
ΔArterial		-1.16[a]					
		(-7.45)					
ΔGallons			-0.21[a]				
			(-4.02)				
ΔDelay				-0.49[a]			
				(-6.01)			
ΔTTI					-0.52[a]		
					(-5.92)		
ΔStress						-0.63[a]	
						(-6.41)	
ΔCongestion							-0.03[b]
							(-2.30)
Adj. R^2	40.4%	40.7%	37.9%	39.4%	39.3%	39.7%	37.0%

Table 31: Directional Index with Alternative Commuting Cost Measures and Workforce Demographics

OLS regression results for urban areas for which Texas A&M Transportation Institute commuting cost proxies were available. The independent variables are changes in median income ($\Delta MedInc$), rental value of agricultural land per acre ($\Delta AgValue$), the percent of civilian labor force that is ($\Delta Female$), freeway miles driven per auto commuter ($\Delta Freeway$), arterial street miles driven per auto commuter ($\Delta Arterial$), annual excess fuel consumed due to traffic congestion per auto commuter ($\Delta Gallons$), vehicle hours of delay due to traffic congestion per auto commuter ($\Delta Delay$), travel time index as the ratio of peak travel time to free-flow travel time as measured in both peak and non-peak travel directions (ΔTTI), commuter stress as measured by the ratio of peak travel time to free-flow travel time in the peak travel direction during the peak travel period ($\Delta Stress$), and annual vehicle and time congestion cost in dollars ($\Delta Congestion$). Area ($Area$) is the 2000 urban area spatial size in square miles. t-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% level sis indicated by a, b, or c superscripts, respectively.

Dependent Variable: Directional Index

Variable	N = 803						
Intercept	88.09[a]	73.24[a]	80.99[a]	84.36[a]	84.95[a]	84.53[a]	85.09[a]
	(34.11)	(25.94)	(24.19)	(29.58)	(29.33)	(30.00)	(30.90)
$\Delta income * 10^{-3}$	1.10[a]	1.22[a]	1.19[a]	1.18[a]	1.18[a]	1.18[a]	1.17[a]
	(7.00)	(7.77)	(7.22)	(7.16)	(7.18)	(7.13)	(7.14)
$\Delta AgValue * 10^{-6}$	96.21	114.60	12.08	7.29	6.36	7.04	-13.65
	(0.91)	(1.07)	(0.11)	(0.07)	(0.06)	(0.06)	(-0.12)
Area	-0.13[a]	-0.12[a]	-0.12[a]	-0.12[a]	-0.12[a]	-0.12[a]	-0.12[a]
	(-23.01)	(-22.49)	(-19.61)	(-18.92)	(-19.29)	(-18.51)	(-20.56)
$\Delta Female$	-0.59[a]	-0.49[a]	-0.16	-0.23	-0.25	-0.23	-0.27
	(-3.23)	(-2.71)	(-0.85)	(-1.20)	(-1.32)	(-1.23)	(-1.43)
$\Delta Freeway$	-1.34[a]						
	(-8.98)						
$\Delta Arterial$		-1.31[a]					
		(-8.79)					
$\Delta Gallons$			0.08				
			(1.46)				
$\Delta Delay$				-0.03			
				(-0.40)			
ΔTTI					-0.08		
					(-0.87)		
$\Delta Stress$						-0.07	
						(-0.62)	
$\Delta Congestion$							-0.02
							(-1.60)
Adj. R^2	46.5%	46.3%	41.2%	41.1%	41.1%	41.1%	41.3%

Table 32: Population Density Gradient with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas (Year 2000)

OLS regression results of the population density gradient with the unconstrained sample, less the five most populous urban areas. For each urban area, the independent variables are median income (*Medinc*), population (*Pop*), the rental value of agricultural land per acer (*AgValue*), the percent of households with at least one vehicle (*Vehicles*), the percent of commuters using public transit (*Transit*), and the percent of the civilian labor force that is female (*Female*). For those urban areas for which Texas A&M Transportation Institute commuting cost proxies were available, the additional independent variables are freeway miles driven per auto commuter (*Freeway*), arterial street miles driven per auto commuter (*Arterial*), annual excess fuel consumed due to traffic congestion per auto commuter (*Gallons*), vehicle hours of delay due to traffic congestion per auto commuter (*Delay*), travel time index (*TTI*), commuter stress index (*Stress*), and annual vehicle and time congestion cost in dollars (*Congestion*). *t*-statistics are listed in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c, respectively.

Dependent Variable: Population Density Gradient (2000)

Variable	N = 435		N = 92						
Intercept	0.06	0.30	0.12	0.17	0.18	0.17	0.11	-0.12	0.16
	(0.07)	(1.32)	(0.66)	(0.88)	(0.97)	(0.92)	(0.28)	(-0.39)	(0.89)
MedInc$*10^{-6}$	-9.14[a]	-9.58[a]	-6.52[b]	-6.98[b]	-6.99[b]	-6.72[b]	-7.19[b]	-8.65[b]	-6.77[b]
	(-2.62)	(-3.10)	(-2.10)	(-2.23)	(-2.23)	(-2.11)	(-2.18)	(-2.53)	(-2.16)
Pop$*10^{-7}$	-6.97[a]	-7.69[a]	-1.80[a]	-1.86[a]	-1.87[a]	-1.67[a]	-1.91[a]	-2.07[a]	-1.54[b]
	(-7.02)	(-7.40)	(-3.92)	(-4.02)	(-3.36)	(-2.74)	(-3.53)	(-4.20)	(-2.55)
AgValue$*10^{-7}$	25.10	-16.00	0.54	-4.75	-5.59	-5.62	-5.73	-9.11	-2.33
	(0.32)	(-0.20)	(0.01)	(-0.12)	(-0.14)	(-0.14)	(-0.14)	(-0.23)	(-0.06)
Female$*10^{-3}$	13.19[a]	12.41[a]	8.73[b]	6.97[c]	6.94[c]	7.41[c]	7.03[c]	8.08[b]	7.76[c]
	(3.01)	(2.83)	(2.22)	(1.82)	(1.75)	(1.90)	(1.84)	(2.08)	(1.99)

Table 32: Population Density Gradient with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas (Year 2000)

*Vehicles*10^-3^*	2.07								
	(0.22)								
Transit		0.02[c]							
		(1.73)							
*Freeway*10^-3^*			-3.90						
			(-1.50)						
*Arterial*10^-4^*				3.35					
				(0.10)					
*Gallons*10^-4^*					2.31				
					(0.05)				
*Delay*10^-3^*						-1.05			
						(-0.49)			
TTI							0.06		
							(0.19)		
Stress								0.26	
								(1.16)	
*Congestion*10^-5^*									-6.56
									(-0.82)
Adj. R^2	14.4%	15.0%	28.7%	26.8%	26.8%	27.0%	26.8%	27.9%	27.4%

Table 33: Population Density Gradient with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas (Year 2010)

OLS regression results of the population density gradient with the unconstrained sample, less the five most populous urban areas. For each urban area, the independent variables are median income ($Medinc$), population (Pop), the rental value of agricultural land per acer ($AgValue$), the percent of households with at least one vehicle ($Vehicles$), the percent of commuters using public transit ($Transit$), and the percent of the civilian labor force that is female ($Female$). For those urban areas for which Texas A&M Transportation Institute commuting cost proxies were available, the additional independent variables are freeway miles driven per auto commuter ($Freeway$), arterial street miles driven per auto commuter ($Arterial$), annual excess fuel consumed due to traffic congestion per auto commuter ($Gallons$), vehicle hours of delay due to traffic congestion per auto commuter ($Delay$), travel time index (TTI), commuter stress index ($Stress$), and annual vehicle and time congestion cost in dollars ($Congestion$). t-statistics are listed in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by a, b, or c, respectively.

Dependent Variable: Population Density Gradient (2010)

Variable	N = 303		N = 77						
Intercept	-0.81	0.51[b]	0.35[b]	0.33[b]	0.35[b]	0.36[b]	0.52	0.34	0.36[b]
	(-0.64)	(2.26)	(2.43)	(2.08)	(2.48)	(2.51)	(1.59)	(1.47)	(2.52)
$MedInc*10^{-6}$	-4.83	-3.32	-4.26[c]	-3.52[c]	-3.78[c]	-3.76[c]	-3.13	-3.71	-3.48
	(-1.46)	(-1.15)	(-1.91)	(-1.68)	(-1.79)	(-1.71)	(-1.38)	(-1.62)	(-1.60)
$Pop*10^{-8}$	-19.53[a]	-20.84[a]	-5.27[a]	-5.03[a]	-5.46[a]	-5.32[a]	-4.65[a]	-5.16[a]	-4.92[a]
	(-4.98)	(-5.07)	(-3.52)	(-3.37)	(-3.33)	(-3.00)	(-2.75)	(-3.28)	(-2.78)
$AgValue*10^{-7}$	29.10	-2.75	13.30	12.20	12.60	13.10	12.70	12.90	13.00
	(0.46)	(-0.04)	(0.57)	(0.52)	(0.54)	(0.56)	(0.54)	(0.55)	(0.56)
$Female*10^{-4}$	34.90	35.20	7.82	6.39	4.13	5.92	9.67	5.92	7.59
	-1.01	(1.02)	(0.37)	(0.30)	(0.19)	(0.27)	(0.44)	(0.26)	(0.35)

Table 33: Population Density Gradient with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas (Year 2010)

Vehicles	0.01								
	(1.06)								
*Transit*10⁻³*		9.96							
		(0.56)							
*Freeway*10⁻³*			2.23						
			(0.81)						
*Arterial*10⁻³*				1.45					
				(0.41)					
*Gallons*10⁻³*					1.94				
					(0.52)				
*Delay*10⁻⁴*						4.38			
						(0.23)			
TTI							-0.17		
							(-0.54)		
Stress								0.03	
								(0.12)	
*Congestion*10⁻⁵*									-1.48
									(-0.18)
Adj. R^2	9.9%	9.6%	20.2%	19.7%	19.8%	19.6%	19.8%	19.5%	19.5%

Table 34: Directional Index with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas

OLS regression results of the directional index (DX) with the unconstrained sample, less the five most populous urban areas. The independent variables are the changes in median income ($\Delta income$), the percent of households with at least one vehicle ($\Delta Vehicles$), the percent of commuters using public transportation, excluding taxis ($\Delta Transit$), the rental value of agricultural land ($\Delta AgValue$), and the percent of civilian labor force that is (Female). For those urban areas for which Texas A&M Transportation Institute commuting costs *proxies* were available, the additional independent variables are freeway miles drive per auto commuter ($Freeway$), arterial street miles driven per auto commuter ($Arterial$), annual excess fuel consumed due to traffic congestion per auto commuter ($Gallons$), vehicle hours of delay due to traffic congestion per auto commuter ($Delay$), travel time index (TTI), commuter stress index ($Stress$), and annual vehicle and time congestion costs in dollar ($Congestion$). t-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% levels is denoted by a, b or c, respectively.

Dependent Variable: Directional Index

Variable	N = 2,867		N = 762						
Intercept	36.21[a]	36.12[a]	90.05[a]	76.37[a]	94.67[a]	95.57[a]	95.66[a]	96.73[a]	91.01[a]
	(19.08)	(19.15)	(35.14)	(27.94)	(25.50)	(32.72)	(32.57)	(33.63)	(33.46)
$\Delta income*10^{-4}$	11.10[a]	7.11[a]	9.83[a]	10.40[a]	10.10[a]	10.50[a]	10.70[a]	10.10[a]	10.00[a]
	(8.98)	(6.06)	(6.55)	(7.20)	(6.57)	(6.97)	(7.08)	(6.71)	(6.57)
$\Delta AgValue*10^{-4}$	5.71[a]	4.18[a]	0.67	1.22	-0.20	-0.28	-0.28	-0.37	-4.33
	(3.58)	(2.60)	(0.66)	(1.25)	(-0.20)	(-0.28)	(-0.27)	(-0.37)	(-0.42)
$Pop*10^{-4}$	-1.48[a]	-1.42[a]	-1.41[a]	-1.32[a]	-1.45[a]	-1.42[a]	-1.42[a]	-1.42[a]	-1.45[a]
	(-22.25)	(-21.25)	(-9.63)	(-9.32)	(-9.64)	(-9.61)	(-9.59)	(-9.72)	(-9.71)
$\Delta Female$	-0.07	-0.11	-0.52[a]	-0.57[a]	-0.32[c]	-0.50[a]	-0.50[a]	-0.42[b]	-0.36[c]
	(-0.48)	(-0.76)	(-2.89)	(-3.31)	(-1.76)	(-2.71)	(-2.72)	(-2.36)	(-1.96)

Table 34: Directional Index with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas

ΔVehicles	-1.26[a]								
	(-2.79)								
ΔTransit		3.73[a]							
		(6.29)							
ΔFreeway			-0.93[a]						
			(-6.23)						
ΔArterial				-1.36[a]					
				(-9.93)					
ΔGallons					-0.15[b]				
					(-2.41)				
ΔDelay						-0.50[a]			
						(-5.22)			
ΔTTI							-0.51[a]		
							(-5.16)		
ΔStress								-0.77[a]	
								(-6.31)	
ΔCongestion									-0.04[a]
									(-3.31)
Adj. R^2	22.5%	23.3%	20.0%	25.7%	16.6%	18.9%	18.8%	20.2%	17.2%

Table 35: Directional Index with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas

OLS regression results of the directional index (DX) with the unconstrained sample, less the five most populous urban areas. The independent variables are the changes in median income ($\Delta income$), the percent of households with at lease one vehicle ($\Delta Vehicles$), the percent of commuters using public transportation, excluding taxis ($\Delta Transit$), the rental value of agricultural land ($\Delta AgValue$), and the percent of civilian labor force that is ($Female$). For those urban areas for which Texas A&M Transportation Institute commuting costs proxies were available, the additional independent variables are freeway miles drive per auto commuter ($Freeway$), arterial street miles driven per auto commuter ($Arterial$), annual excess fuel consumed due to traffic congestion per auto commuter ($Gallons$), vehicle hours of delay due to traffic congestion per auto commuter ($Delay$), travel time index (TTI), commuter stress index ($Stress$), and annual vehicle and time congestion costs in dollar ($Congestion$). t-statistics are in parentheses. Statistical significance at the 1%, 5%, and 10% levels is denoted by a, b or c, respectively.

Dependent Variable: Directional Index

Variable	N = 2,867		N = 762						
$Intercept$	38.54[a]	38.43[a]	88.72[a]	74.81[a]	92.73[a]	95.11[a]	95.26[a]	96.34[a]	89.15[a]
	(19.81)	(19.87)	(32.95)	(26.22)	(23.99)	(30.90)	(30.77)	(31.79)	(31.35)
$\Delta income*10^{-4}$	11.60[a]	7.41[a]	10.10[a]	10.60[a]	10.30[a]	10.80[a]	11.00[a]	10.30[a]	10.30[a]
	(9.42)	(6.32)	(6.55)	(7.19)	(6.58)	(6.98)	(7.10)	(6.70)	(6.58)
$\Delta AgValue*10^{-4}$	5.08[a]	3.50[b]	0.70	1.28	-0.15	-0.29	-0.29	-0.39	-0.33
	(3.18)	(2.18)	(0.68)	(1.28)	(-0.15)	(-0.28)	(-0.28)	(-0.38)	(-0.31)
$Area$	-0.15[a]	-0.14[a]	-0.11[a]	-0.10[a]	-0.11[a]	-0.12[a]	-0.11[a]	-0.12[a]	-0.11[a]
	(-22.35)	(-21.44)	(-7.43)	(-7.14)	(-7.45)	(-7.69)	(-7.69)	(-7.84)	(-7.41)
$\Delta Female$	-0.10	-0.15	-0.51[a]	-0.57[a]	-0.30	-0.48[b]	-0.50[b]	-0.40[b]	-0.33[c]
	(-0.70)	(-0.99)	(-2.69)	(-3.15)	(-1.56)	(-2.56)	(-2.58)	(-2.16)	(-1.73)

Table 35: Directional Index with Alternative Commuting Cost Measures and Workforce Demographics, Excluding the 5 Largest Urban Areas

ΔVehicles	-1.32[a]								
	(-2.93)								
ΔTransit		3.88[a]							
		(6.58)							
ΔFreeway			-0.93[a]						
			(-6.10)						
ΔArterial				-1.39[a]					
				(-9.91)					
ΔGallons					-0.13[b]				
					(-2.13)				
ΔDelay						-0.54[a]			
						(-5.47)			
ΔTTI							-0.54[a]		
							(-5.43)		
ΔStress								-0.82[a]	
								(-6.54)	
ΔCongestion									-0.03[a]
									(-2.77)
Adj. R^2	22.6%	23.5%	16.4%	22.4%	12.8%	15.6%	15.6%	17.0%	13.2%

APPENDIX: URBAN SPRAWL COMPARATIVE STATICS

1. Introduction

The comparative statics referred to in Equation 3 are derived in this appendix. I will closely follow Turnbull (n.d.), with minor deviations in notation, and exploit the duality between consumer utility maximization and expenditure minimization to derive the comparative statics. There are two categories of cities used in urban spatial structure analysis, *closed* and *open*, and both characterize the trade-off consumers face between the costs of transportation and housing. Costless migration to and from an urban area is possible in the open city, with endogenously determined population being the adjustment mechanism that ensures the same level of utility both inside or outside the urban area. Conversely, in the closed city where costless migration is not possible, population is exogenous and utility and the spatial area of the city are determined within the system. Parameters that are common to both closed and open cities are income, transportation costs, and housing rents. To remain consistent with the rest of this paper by following Brueckner and Fansler (1983), I derive comparative statics for the closed city model only.

2. Closed City Model

Urban residents (consumers) commute to a central business district (CBD) at a round-trip cost represented by the function $t(x)$, where x is measured in miles from the CBD, and earn identical income, y.[32] The consumers' utility, $u(\cdot)$, is expressed as

$$u(c, q) = u, \tag{23}$$

where u is the highest level of utility attainable across the urban area, and utility is a regular strictly quasi-concave function of a composite non-housing good, c, and a housing good, q, measured in

[32] Assume the CBD is located at $x = 0$.

square feet of floor space. The composite good is sufficient to represent all consumption other than housing, including household saving. The assumed regular strict quasi-concavity of $u(\cdot)$ has three important implications for the present analysis. The first is that the function is increasing in its arguments so that marginal utilities of both goods are positive. The second is that the bordered Hessian matrix of second derivatives is negative semi-definite, which ensures that the second-order sufficient conditions are met.[33] And the third is that there is a diminishing marginal rate of substitution between c and q. Moreover, strict quasi-concavity necessarily implies that consumers' indifference curves between c and q are strictly convex.[34]

Consumer utility theory provides two methods of solving the consumer's problem of choosing the optimal consumption bundle: the direct and duality methods. In the direct method, consumers wish to maximize utility from consumption given available income. Conversely, duality in consumer theory provides the indirect method as an alternative for determining the optimal consumption bundle for consumers by minimizing the expenditures necessary to reach a given utility level. Here, I use the indirect method. Under the assumption that all urban consumers achieve the same level of utility, u, variation in land rent, $r(x)$, charged by landlords across the urban area ensures that this equilibrium holds, as a resident who lives farther from the CBD will be compensated for increased commuting costs with lower housing rent. The price of the non-housing good is the same everywhere in the city.[35] Thus, consumers face the following budget constraint that all expenditures must equal all available income:

[33] The determinant of the Hessian matrix of a strictly quasi-concave function is positive, but it can be zero at isolated points. With the assumed regular strictly quasi-concave function, the possibility of such points in eliminated.

[34] The importance of strict quasi-concavity of the consumers' utility function is that the axioms of completeness, continuity, strict monotonicity, and strict convexity of consumers' preferences over c and q hold.

[35] The price of c is assumed to be 1 for the present discussion. This assumption does not qualitatively alter the results.

$$y - t(x) - c - r(x)q = 0 \tag{24}$$

Expressed in dual-form, the consumer's problem becomes one of expenditure minimization, and its solutions represent the income-compensated demand functions that minimize expenditures, subject to the constraint that all consumers attain utility level u.[36]

$$\{c(r,u), q(r,u)\} \equiv arg\,min\, r(x)q + c \;\; s.t. \; u = u(c,q). \tag{25}$$

Since the consumer expenditure function, $e(r(x),u)$, defines the amount of money needed to attain utility level u at distance x, the identity $e(r(x),u) = y - t(x)$ follows. Further, satisfaction of Equation 25 is ensured by construction of c. Defined implicitly, the expenditure function becomes $F(r,x,u,y) = e(r,u) - y + t(x) = 0$. By Shephard's lemma, the demand for housing is the derivative of the expenditure function with respect to r, or $\frac{\partial F}{\partial r} = e_r = q$.[37] As $q > 0$, the derivative of the implicit expenditure function with respect to r does not vanish, allowing us to invoke the implicit function theorem to solve the identity,

$$e(r(x),u) \equiv y - t(x), \tag{26}$$

for a differentiable function, $r(x)$. Differentiating Equation 26 with respect to x yields

$$e_r(r(x),u)r_x(x) = -t_x(x).^{38} \tag{27}$$

[36] Income-compensated demand functions are also known as Hicksian demand functions.
[37] Shephard's lemma states that the demand functions that minimize expenditures are the derivatives of the expenditure function with respect to the price of each good.
[38] Subscripts denote partial derivatives throughout this appendix.

Recalling that $e_r = q$ from above and substituting the solution obtained from the implicit function theorem, $q(r(x), u)$, into Equation 26 yields the rent gradient as the first derivative of the rent function with respect to distance,

$$r_x(x) = -\frac{t_x(x)}{q(r(x),u)} < 0, \tag{28}$$

which suggests that household rent is a declining function of household distance from the CBD. Figure 10 illustrates this central tenet of the model: that urban residents face a trade-off between commuting costs and rent.

Moreover, the convex rent function in Figure 10 can ensured by a positive second derivative of the rent function with respect to distance. Differentiating Equation 29 with respect to x provides

$$r_{xx}(x) = -\frac{t_{xx}(x)q(r(x),u)+q_r(r(x),u)r_x(x)t_x(x)}{q(r(x),u)^2}. \tag{29}$$

In order to determine if r_{xx} is indeed positive, we must first determine the signs of q_r and t_{xx}. That $q_r < 0$ follows from standard consumer theory that the demand for a good is a decreasing function of the good's own-price. Further, as long as the cost of each additional mile of commute is non-increasing, $t_{xx} \leq 0$, combined with the previous result of $r_x < 0$, enables us to deduce that $r_{xx} > 0$.

To continue developing the model, we must also consider the process by which urban landlords obtain land to develop and establish the land market equilibrium conditions. In the stylized circular urban area where urban landlords compete for land with rural agricultural users, landlords must outbid agricultural users for land at the urban-rural boundary, $\bar{x}$.[39] The profit maximizing landlord auctions off his land to the highest bidder, extracting the maximum rent from each resident at all x, and thus operates as a monopsonist over his unique parcel(s) of land. As landlords are profit maximizing, equilibrium is obtained at the zero-profit condition Equation 31, where urban land rents equal agricultural land rents, denoted r_a.

$$r(\bar{x}) = r_a \tag{30}$$

It then follows that inside the urban area at all distances $x < \bar{x}$, $r > r_a$, and at all distances $x > \bar{x}$, $r < r_a$. The second equilibrium condition constrains the urban population, L, to an area no larger than a circle with radius $\bar{x}$,

$$2\pi \int_0^{\bar{x}} \frac{x}{q(r(x),u)} dx = L. \tag{31}$$

Consistent with the empirical method of this paper, equilibrium condition Equation 31 measures the urban population by treating the urban area as a series of adjacent rings, each of width x, emanating from the CBD.

[39] The urban-rural boundary separates the urbanized area from its surrounding area.

Figure 11 illustrates the equilibrium established by Equation 31. At all distances x to the left of $\bar{x}$, residential rents are bound between $[r_a, r(0)]$.

3. Comparative Statics

For transportation cost comparative static derivation, Wheaton (1974) assumed a linear cost function. We will remain consistent with his original paper, such that $t(x) \stackrel{\text{def}}{=} tx$, resulting in transportation costs entering the rent function as a parameter; $r(x; t, y, u)$ when deriving commuting cost comparative statics. The first step in deriving the comparative static predictions of the model is to obtain the Jacobian matrix of first derivatives. Recalling that in the closed city our solutions provide the equilibrium urban spatial area and utility, and that Equations 26, 30, and 31 fully describe the land market, substitute Equation 30 into Equation 26 *at the urban-rural boundary*, $\bar{x}$, to reduce the model to the following system of two equations:

$$e(r_a, u) = y - t(\bar{x}) \qquad (32)$$

$$2\pi \int_0^{\bar{x}} \frac{x}{q(r(x), u)} dx = L \qquad (33)$$

The system immediately above yields the following Jacobian, denoted $\boldsymbol{J}$,:

$$\boldsymbol{J} = \begin{bmatrix} t_x(\bar{x}^*) & e_u(r_a, u^*) \\ \frac{2\pi\bar{x}^*}{q(r_a, u^*)} & -2\pi \int_0^{\bar{x}^*} \left(\frac{x(q_r r_u + q_u)}{q(r, u^*)^2} \right) dx \end{bmatrix} \qquad (34)$$

and its determinant,

$$|\boldsymbol{J}| = t_x(\bar{x}^*) \left(-2\pi \int_0^{\bar{x}^*} \left(\frac{x(q_r r_u + q_u)}{q(r, u^*)^2} \right) dx \right) - e_u(r_a, u^*) \frac{2\pi\bar{x}^*}{q(r, u^*)}. \qquad (35)$$

Factoring out -2π and rearranging Equation 35 yields

$$|\boldsymbol{J}| = -2\pi \left[t_x(\bar{x}^*) \int_0^{\bar{x}^*} \left(\frac{x(q_r r_u + q_u)}{q(r,u^*)^2} \right) dx + \frac{\bar{x}^* e_u(r_a,u^*)}{q(r,u^*)} \right]. \tag{36}$$

To invoke the implicit function theorem to solve Equations 32 and 33 for $\bar{x}^*$ and u^*, $|\boldsymbol{J}| \neq 0$. We can conclude that $|\boldsymbol{J}| < 0$ from examination of both bracketed terms. The notion that positive urban spatial area (the integral in Equation 33 is positive) necessarily leads to positive marginal commuting costs ($t_x > 0$ at the boundary) from the urban periphery, causes the first bracketed term to be positive. As $\left(\bar{x}^*, q(r,u^*)\right) > (0,0)$, and $e_u(r_a,u^*) > 0$ as expenditure must rise if the consumer attains a higher indifference curve, the second bracketed term is likewise zero. Thus, we use the implicit function theorem to solve for $\bar{x}^*(y,L,r_a,t)$ and $u^*(y,L,r_a,t)$. Total differentiation of the system yields:

$$\boldsymbol{J} \begin{bmatrix} d\bar{x}^* \\ du^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & -q(r_a,u^*) & -\bar{x} \\ \int_0^{\bar{x}^*} \frac{2\pi x q_r r_y}{q^2} dx & 1 & 0 & \int_0^{\bar{x}^*} \frac{2\pi x q_r r_t}{q^2} dx \end{bmatrix} \begin{bmatrix} dy \\ dL \\ dr_a \\ dt \end{bmatrix} \tag{37}$$

We employ Cramer's rule below to solve Equation 37 for the comparative statics $\partial\bar{x}^*/\partial y$, $\partial\bar{x}^*/\partial L$, $\partial\bar{x}^*/\partial r_a$, and $\partial\bar{x}^*/\partial t$.

3.1 Consumer Incomes

As consumer incomes rise, urban residents are able to attain a higher level of utility. Thus, the model predicts that residents will prefer more housing space as incomes rise, moving outward from the CBD and expanding the urban area footprint. Solving Equation 37 for the partial derivative of $\bar{x}$ with respect to y yields

$$\partial\bar{x}^*/\partial y = \frac{\begin{vmatrix} 1 & e_u(r_a,u^*) \\ \int_0^{\bar{x}^*} \frac{2\pi x q_r r_y}{q^2} dx & -2\pi \int_0^{\bar{x}^*} \left(\frac{x(q_r r_u + q_u)}{q^2} \right) dx \end{vmatrix}}{|\boldsymbol{J}|}, \tag{38}$$

which reduces to

$$\partial \bar{x}^* / \partial y = -\frac{2\pi}{|J|}\left(\int_0^{\bar{x}^*}\left(\frac{x(q_r r_u + q_u)}{q^2}\right)dx + e_u(r_a, u^*)\int_0^{\bar{x}^*}\frac{x q_r r_y}{q^2}dx\right) \tag{39}$$

To determine the sign of $\partial \bar{x}^* / \partial y$, we must examine the signs of the terms in parentheses, but as the first term is positive and the second is negative, the sign is ambiguous at this point. Substituting the land rent function, $r(x; t, y, u)$, into Equation 26 yields the following identity:

$$e(\,r(x; t, y, u), u) \equiv y - tx \tag{40}$$

Differentiating Equation 40 provides the following derivatives:[40]

$$e_r r_y = 1 \quad \Rightarrow \quad r_y = \frac{1}{q} > 0 \tag{41}$$

$$e_r r_u + e_u = 0 \quad \Rightarrow \quad r_u = -\frac{e_u}{q} < 0 \tag{42}$$

$$e_r r_t = -x \quad \Rightarrow \quad r_t = -\frac{x}{q} < 0. \tag{43}$$

Substituting r_y and r_u into Equation 39 yields.

$$\partial \bar{x}^* / \partial y = -\frac{2\pi}{|J|}\left(\int_0^{\bar{x}} -\frac{x}{q^3}(q_r e_u(r, u) + q_u)dx + \int_0^{\bar{x}}\frac{x}{q^3}q_r dx e_u(r_a, u)\right) \tag{44}$$

Noting that $de_u/dx = e_{ur} r_x = q_u r_x < 0$ as rent declines with increased distance from the CBD, shows that e_u is declining in x. Therefore, after combining integrals below, $q_r\big(e_u(r_a, u^*) - e_u(r, u^*)\big) > 0$, as $e_u(r_a, u^*) - e_u(r, u^*) < 0$ at all distances within the urban-

[40] Recall that $e_r = q$ by Shephard's lemma.

rural boundary. With housing as a normal good, $qq_u > 0$ allows us to determine the sign of $\partial\bar{x}^*/_{\partial y}$, and we conclude that increases in residents' incomes causes the city size to expand, or

$$\partial\bar{x}^*/_{\partial y} = -\frac{2\pi}{|J|}\left(\int_0^{\bar{x}^*}\frac{x}{q^3}\left(q_r\left(e_u(r_a,u^*) - e_u(r,u^*)\right) + qq_u\right)dx\right) > 0. \tag{45}$$

3.2 Urban Population

Along with increases income, it follows from Equation 33 that increases in population will increase the spatial area of the city, as well. Solving Equation 37 for $\partial\bar{x}^*/_{\partial L}$ yields

$$\partial\bar{x}^*/_{\partial L} = \frac{\begin{vmatrix} 0 & e_u(r_a,u^*) \\ 1 & -2\pi\int_0^{\bar{x}^*}\left(\frac{x(q_r r_u + q_u)}{q^2}\right)dx \end{vmatrix}}{|J|} = -\frac{e_u(r_a,u^*)}{|J|} > 0 \tag{46}$$

As shown above when determining the sign of Equation 36, $e_u > 0$ and $|J| < 0$. Therefore, $\partial\bar{x}^*/_{\partial L} > 0$.

3.3 Agricultural Land Rent

As agricultural land rents increase, land developers face greater competition and will substitute taller building heights toward the city center for larger residential dwellings. The model predicts that this substitution will cause the urban area footprint to contract. We solve Equation 38 for the partial derivative of $\partial\bar{x}^*/_{\partial r_a}$ to obtain

$$\partial\bar{x}^*/_{\partial r_a} = \frac{\begin{vmatrix} -q(r_a,u^*) & e_u(r_a,u^*) \\ 0 & -2\pi\int_0^{\bar{x}^*}\left(\frac{x(q_r r_u + q_u)}{q^2}\right)dx \end{vmatrix}}{|J|}, \tag{47}$$

which, with rearranging, becomes

$$\partial\bar{x}^*/_{\partial r_a} = 2\pi \frac{q(r_a,u^*)}{|J|} \int_0^{\bar{x}^*} \left(\frac{x(q_r r_u + q_u)}{q^2}\right) dx < 0. \tag{48}$$

An increase in agricultural rents leads to a reduction in the urban spatial area as both the integrand and numerator outside the integral are positive and $|\boldsymbol{J}| < 0$.

3.4 Transportation Costs

As with agricultural rents, the model predicts that increases in transportation costs induce residents to favor smaller dwellings closer to the CBD over larger dwellings and longer commute times. Finally, solving Equation 38 for the partial derivative of spatial area with respect to transportation costs yields

$$\partial\bar{x}^*/_{\partial t} = \frac{\begin{vmatrix} -\bar{x} & e_u(r_a,u^*) \\ \int_0^{\bar{x}^*} \frac{2\pi x q_r r_t}{q^2} dx & -2\pi \int_0^{\bar{x}^*} \left(\frac{x(q_r r_u + q_u)}{q^2}\right) dx \end{vmatrix}}{|J|}. \tag{49}$$

After rearranging, Equation 49 becomes

$$\partial\bar{x}^*/_{\partial t} = -\frac{2\pi\bar{x}}{|J|} \left(\int_0^{\bar{x}^*} \left(\frac{x(q_r r_u + q_u)}{q^2}\right) dx + \int_0^{\bar{x}^*} \left(\frac{q_r r_t}{q^2}\right) dx \, e_u(r_a, u^*) \right). \tag{50}$$

To determine the sign of $\partial\bar{x}^*/_{\partial t}$, we follow a similar procedure to that which we used for $\partial\bar{x}^*/_{\partial y}$. Substituting r_t and r_u into Equation 50 yields

$$\partial\bar{x}^*/_{\partial t} = -\frac{2\pi\bar{x}}{|J|} \left(\int_0^{\bar{x}^*} \left(\frac{x(q_r e_u(r,u) + q_u)}{q^3}\right) dx - \int_0^{\bar{x}^*} \left(\frac{q_r x}{q^3}\right) dx \, e_u(r_a, u^*) \right), \tag{51}$$

and rearranging provides

$$\partial\bar{x}^*/_{\partial t} = -\frac{2\pi\bar{x}}{|J|} \left(\int_0^{\bar{x}^*} \frac{x}{q^3} \Big(q_r \big(e_u(r_a, u^*) - e_u(r, u^*) \big) \Big) dx + q_u \right) dx < 0. \tag{52}$$

Twice differentiating the expenditure function with respect to u and t yields $de_u/dt = e_{ut}r_t = q_u r_t < 0$. Because rent declines with increased commuting costs to the CBD, we determine that that e_u is declining in t, which allows us to determine the sign in Equation 52, or that ${\partial \bar{x}^*}/{\partial t} < 0$.

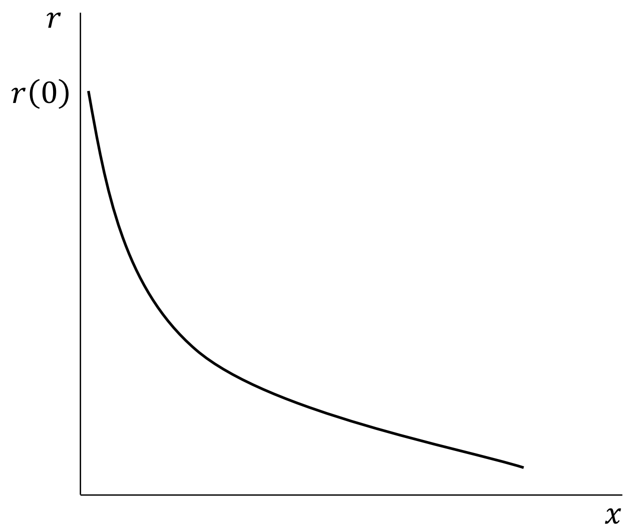

Figure 10: Graph of the Rent Function

This figure graphs rent as a function of distance. Rent, r, and distance, x, are represented on the vertical and horizontal axes, respectively.

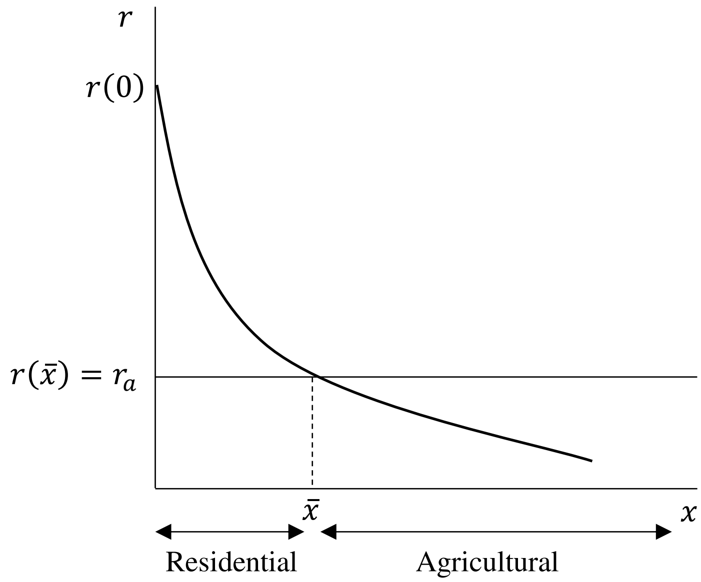

Figure 11: Graph of the Rent Function

This figure graphs rent as a function of distance. Rent, r, and distance, x, are represented on the vertical and horizontal axes, respectively. The urban-rural boundary is denoted $\bar{x}$.