UNIVERSITY OF CALIFORNIA

Los Angeles

Three Essays in Macro-Finance.

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

David Ciaran Lindsay

2022

ABSTRACT OF THE DISSERTATION

Three Essays in Macro-Finance.

by

David Ciaran Lindsay
Doctor of Philosophy in Economics
University of California, Los Angeles, 2022
Professor Andrew G. Atkeson, Co-Chair
Professor Pierre-Olivier Weill, Co-Chair

In Chapter 1, I use a structural approach, to quantify the effect of land-use regulations on different age and education groups. I estimate a dynamic spatial structural equilibrium model of household location choice, local housing supply, and amenity supply. I show that in the long-run, removing land-use restrictions benefits all household groups and increases aggregate consumption by 7.1%. These consumption gains vary across households, less educated and younger households see increases in consumption about twice as large as more educated or older households. In contrast, in the short-run, removing land-use regulations reduces the consumption of older-richer homeowners while increasing the consumption of younger renters. In a counterfactual 1990-2019 transition, abolishing land-use regulations reduces the consumption of households born before the mid-1960s, while increasing consumption of more recent generations.

In Chapter 2, co-authored with Mahyar Kargar, Benjamin Lester, Shuo Liu, Pierre-Olivier Weill, Diego Zúñiga, we study liquidity conditions in the corporate bond market during the COVID-19 pandemic. We document that the cost of trading immediately via risky-principal trades dramatically increased at the height of the sell-off, forcing customers to shift toward slower agency trades. Exploiting eligibility requirements, we show that the Federal Reserve's corporate credit facilities have had a positive effect on market liquidity. A structural estimation reveals that customers' willingness to pay for immediacy increased by about 200 bps per dollar of transaction, but quickly subsided after the Fed announced its interventions. Dealers' marginal cost also increased substantially but did not fully subside.

In Chapter 3, co-authored with Diego Zúñiga, we study inter-dealer trading patterns in the US corporate bond market. We document that dealers trade with only a small group of other dealers and that this group of dealers is highly persistent over time. We show that the longer a dealer pair have been trading the more likely that they will continue to trade and the larger the bilateral volume traded between them. We measure trading costs between dealers and show that stronger relationship leads to lower trading costs. Motivated by our empirical work we develop a structural model of trading relationships. The existence of double marginalization leads to inefficiency. We show that the repeated nature of the interactions between dealers allows them to form relationships and hence restore optimality.

The dissertation of David Ciaran Lindsay is approved.

Tyler S. Muir

Matthew Saki Bigio Luks

Pierre-Olivier Weill, Committee Co-Chair

Andrew G. Atkeson, Committee Co-Chair

University of California, Los Angeles

2022

To my friends and family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am deeply indebted to Pierre-Olivier Weill and Andy Atkeson for their invaluable feedback. Their mentorship, support and encouragement were of utmost importance, and it would have been impossible to prepare this thesis without them. I am also eternally grateful to Lee Ohanian for his guidance and support. My work has benefited from suggestions from Saki Bigio, Paula Beltran, Huifeng Chang, Conor Foley, Francois Geerolf, Tomas Guanziroli, Gary Hansen, YongKi Hong, Brian Pustilnik, Diego Zúñiga, and seminar participants at UCLA.

I gratefully acknowledge financial support from the Department of Economics and the Graduate Division. I also want to thank the staff of the Department of Economics at UCLA for their unwavering support, specially Grace Fransisca, Juliana Smith, and Chiara Paz. Finally, I could not have done this without my co-authors, Mahyar Kargar, Benjamin Lester, Shuo Liu, Pierre-Olivier Weill, Diego Zúñiga.

VITA

Education

2017	Master of Arts in Economics, UCLA, Los Angeles, California, USA.
2016	Bachelor of Arts in Mathematics and Economics, Trinity College Dublin, The University of Dublin, Dublin, Ireland.

Experience

2017-2021	Teaching Assistant, Economics Department, UCLA, Los Angeles, California, USA.
2019	Research Assistant, Economics Department, UCLA, Los Angeles, California, USA.
2019-2020	Summer Instructor, Economics Department, UCLA, Los Angeles, California, USA.

Awards

2021-2022	Dissertation Year Fellowship, UCLA, Los Angeles, California, USA.
2018	Ziman Center for Real Estate research Grant, UCLA, Los Angeles, California, USA.

CHAPTER 1

The Heterogeneous Effect of Local Land-Use Restrictions Across US Households

1.1 Introduction

> Exclusionary zoning laws enact barriers to entry that constrain housing supply, which, all else equal, translate into an equilibrium with more expensive housing and fewer homes being built . . . The American Jobs Plan takes important steps to eliminate exclusionary zoning. Specifically, the Unlocking Possibilities Program within the American Jobs Plan is a $5 billion competitive grant program that incentivizes reform of exclusionary zoning . . . [it] incentivizes new land-use and zoning policies to remove those barriers.

[June 2021] Council of Economic Advisors

In recent decades, house prices in the most productive cities, such as New York and the San Francisco Bay Area, have risen substantially, making these cities unaffordable to many low-wage workers. As argued in Glaeser and Gyourko (2018), Gyourko and Molloy (2015), and Glaeser et al. (2005) the literature attributes high housing costs in these cities to local land-use restrictions. Land-use regulations such as height limits, lot size & parking minimums, set-back requirements, environmental reviews, historic preservation, and prolonged approval processes fraught with regulatory discretion, constrain the housing supply and hence raise local housing costs.

These changes to the land-use regulatory environment have led to rising spatial sorting of households along education and city productivity lines (Van Nieuwerburgh and Weill (2010)) as low-wage workers are no longer able to afford housing in high-productivity cities. This increases inequality between workers of different education, as less-educated workers increasingly live in cities that experienced the weakest productivity growth. Furthermore, the tightening of land-use regulations creates inequality between different age cohorts. Older households are usually homeowners and hence are insulated from the increased housing costs and the value of their most important asset rises. In contrast, younger households, who are typically renters, spend an increased fraction of their income on housing and may be unable to afford housing in the more productive cities.

In my paper, I quantify how land-use restrictions affect different age and education cohorts. Owing to the lack of viable natural variability in land-use regulations and general equilibrium effects[1] I build on the seminal work of Roback (1982) and estimate a dynamic spatial equilibrium model of household location choice, local housing supply, and amenity supply. I estimate the local housing supply elasticities and the household's utility parameters. Using the estimated model I show that abolishing land-use regulations increases aggregate output by about 7.1% in the long-run. I show that the benefits from removing land-use restrictions are heterogeneous

[1]The lack of natural experiments is an issue faced by much of this housing literature. With the exception of Diamond et al. (2019), natural experiments are few and far between, and hence the reliance on structural methods.

across age and education groups. The youngest and least educated households see increases in consumption that are about twice as large as the oldest and most educated households. While in the long-run land-use restrictions negatively affect all age and education groups, in the short-run abolishing land-use regulations tends to benefit younger renters at the expense of incumbent older homeowners.

My model features heterogeneous households that differ in terms of their age, homeownership status, education, wealth, and idiosyncratic preference shocks. Each period households choose a city to live in, trading off moving costs, amenities, housing costs, as well as the wage they would receive in each city. Cities with more stringent land-use regulations have lower housing supply elasticities and higher house prices as a result. In equilibrium, housing markets and rental markets clear. The endogenous location choices of households and housing supply functions determine equilibrium house prices and rents.

My model is able to capture two key features of the data. First, there is a positive correlation between city productivity (or amenities) and house prices. This arises as high-productivity or a plentiful supply of amenities induces households to move to the region, increasing house prices. Secondly, there is positive assortative matching wherein more educated households tend to sort into high-productivity high-cost cities, as the rise in their income relative to the rise in housing costs is larger compared to less educated households.

Analyzing policy counterfactuals requires estimates of each city's housing supply elasticity and the household's utility parameters. I estimate housing supply elasticities using international migration shocks as an instrument for housing demand. As has been argued by Saiz (2010), international migration tends to occur along established networks and is therefore orthogonal to changes in city productivity or housing supply shocks. For example, the decision for a Cuban or Mexican to migrate to Los Angeles or Miami reflects the predetermined migration network and economic and political changes in their home country rather than city-specific changes in Los Angeles or Miami. To structurally estimate the household's utility parameters, I exploit the conditional choice structure of the model and maximize the likelihood function using the 5% Census microdata sample.

My paper has four main counterfactual results. In line with Herkenhoff et al. (2017) and Hsieh and Moretti (2019), I show that abolishing land-use regulations increases aggregate output. When land-use regulations are abolished, housing costs are equalized across cities. In the baseline steady-state, prices are higher in more productive regions, so relaxing land-use regulations decreases the relative price of housing in high-productivity cities. Thus, the population of more productive regions expands at the expense of less productive regions, increasing aggregate output. I find that aggregate output rises by approximately 7.1%, when I compare the baseline and counterfactual steady-states. Along with the rise in aggregate output, the utility households receive from amenities also rises by 8.6%, as households tend to relocate to high-productivity and high amenity regions.

Secondly, I show that while all education cohorts witness increased consumption and amenity utility when land-use regulations are relaxed, it is the least educated who see the largest percentage rise in their consumption. When the relative price of housing in high-cost high-productivity cities falls, the consumption of existing low-wage workers rises more, as housing costs consume a larger fraction of their income. Moreover, since affordability relative to income has improved to a greater extent for low-wage workers, they will tend to move to productive cities. Hence, when we compare steady-states, the degree of income dispersion decreases when we relax land-use regulations.

Similarly, I show that when comparing steady-states, younger households see the largest rise in their consumption when land-use regulations are abolished. In my model, all households transition from being a renter when they start their working life to being a homeowner later in life. In the estimated model, wages increase as workers age, reflecting the effect of experience. This combined with a down-payment constraint means that younger workers not only spend a larger fraction of their income renting housing when prices rise, but they also save a larger fraction of their income to meet the down-payment constraint. Quantitatively, I show that the percent rise in consumption is approximately twice as large for young households compared to the oldest cohorts.

Thirdly, in the short-run, I show that land-use regulations can benefit homeowners at the expense of renters. To show this I examine the consumption of different household groups in a transition between steady-states. I create a series for the productivity shocks and amenity supply of each city. The productivity shocks are the Bartik change in city wages using 1990 industry wage bill shares and national changes in wages for each industry, excluding the city itself. Taking these changes as given, I show that transitions with estimated housing supply elasticities tend to benefit existing homeowners at the expense of renters and future generations, as existing homeowners reap large capital gains on their homes, while renters face increasingly costly housing. Given how responsive the political system is to the demands of homeowners, this can perhaps explain the ubiquity of stringent land-use regulations given their seemingly harmful effects.

Motivated by the results from the transitional dynamics as well as how entrenched restrictive land-use regulations are in many cities, I derive my final set of results, examining whether a switch to remote work or creating new cities can replicate the long-run benefits of relaxing land-use regulations. Anecdotally, the rising fraction of remote workers is cited as the cause of the population decline of high-cost cities such as San Francisco and New York during the COVID-19 pandemic. I find that both allowing a fraction of households to work from home and creating new cities increase aggregate output. However, the aggregate rise in output is approximately 1.38% when 20% of workers work remotely or the number of cities increases by 10%. Although I allow all workers to work from home with equal likelihood, the least educated workers see a rise in their consumption about twice as large as the most educated. Similarly, less educated households see consumption growth about three times as large as the most educated households when I increase the number of cities by 10%.

1.2 Related Literature

While my paper makes a number of distinct contributions to the literature, I highlight several aspects of this study that are particularly important. First, my paper quantifies which household groups bear the costs of land-use regulations. I show that in the long-run, younger and less educated households benefit the most (in consumption terms) from removing restrictions. Secondly, my paper shows that in the short-run, land-use regulations creates winners and losers among households. As the economy transitions between its 1990 and 2019 steady-state, restrictive land-use regulations benefit older cohorts at the expense of younger cohorts. Life-time consumption is larger in transitions with the baseline land-use regulations for cohorts born before the mid-1960s, compared to a transition where land-use regulations are abolished. Finally, I show that compared with relaxing land-use restrictions, the effect of remote working or establishing new cities on household consumption is quantitatively small.

This paper is related to the burgeoning literature on housing and macroeconomics, see Piazzesi and Schneider (2016) for an overview, and is particularly related to the macroeconomic effects of land-use regulations. Herkenhoff et al. (2017) and Hsieh and Moretti (2019) examine the impact of land-use regulations have had on aggregate output in the US. Similar to these papers, this study finds that removing land-use restrictions would increase aggregate output. Bunten (2017) shows that the local nature of land-use regulation creates inefficiencies relative to a national planner. Although the increase in aggregate output I observe is smaller than either Herkenhoff et al. (2017) or Hsieh and Moretti (2019), it remains substantial. I find smaller aggregate effects because in my model households are heterogeneous in terms of productivity and so the most productive workers are already located in the most productive cities, limiting the gains from expanding the city. Furthermore, my model features idiosyncratic household preferences which reduces the housing demand elasticity for a city, as not all workers will want to move to a high-productivity city even if it is affordable.

The spatial sorting of households across urban areas is analyzed in Van Nieuwerburgh and Weill (2010), Gyourko et al. (2013), Ganong and Shoag (2017), and Bilal and Rossi-Hansberg (2021). [2]. These papers show that heterogeneous education or

[2]Couture et al. (2020) and Almagro and Domınguez-Iino (2021) examine spatial sorting at the city-level. Couture et al. (2020) shows that in recent decades, an increasing number of wealthy

ability of households leads to the most educated households sorting into so-called "Superstar Cities". In my counterfactual analysis, relaxing land-use restrictions reduces the degree of spatial sorting as lower-wage workers are able to afford these high-productivity regions. Relatedly, Baum-Snow et al. (2018) and Fajgelbaum and Gaubert (2020) examine spatial sorting and its effect on inequality and efficiency.

Another related strand of literature, is the spatial economics literature first formulated in Sherwin (1979) and Roback (1982). My paper shares the dynamic discrete framework of Kennan and Walker (2011), Diamond (2016), Monte et al. (2018), Almagro and Domınguez-Iino (2021), and Schubert (2020). Most closely related to my paper Diamond (2016) further scrutinizes Moretti (2013) result: finding that well-being inequality has increased between college and non-college workers over time due to changes in their location choices. Uniquely my paper incorporates overlapping generations of households who smooth consumption over their life-cycle and differences in housing tenure, allowing an analysis of the impact of regulations across cohorts and owners and renters. Moreover, this study analyzes the out-of-steady dynamics of the spatial economy, which is crucial to determine the short-run impacts of changes to fundamentals across different groups.

A core contribution of my paper is to use the structural model to derive the heterogeneous effects from counterfactual changes in land-use restrictions. Kiyotaki et al. (2011) and Kiyotaki et al. (2020), Favilukis and Van Nieuwerburgh (2017), Favilukis et al. (2019) build structural models to examine how housing policies affect different household groups. The former two papers show that monetary policy shocks

households now live in downtown areas of cities alongside the existing low-income households, creating a U-shaped pattern of sorting in downtown areas. While Almagro and Domınguez-Iino (2021) argues that endogenous amenities play a crucial role in location choices in Amsterdam.

benefits existing homeowners at the expense of current renters while the latter papers examine city-level policies [3]. Unlike these papers my model features a diverse set of cities from which households can choose from each period.

In terms of estimation methodology my paper uses the conditional choice probability (CCP) techniques borrowed from the industrial organization literature (Rust (1987), Hotz and Miller (1993), and Arcidiacono and Miller (2011), Arcidiacono and Ellickson (2011)). These estimation techniques have spatial featured in Scott (2014), Almagro and Domınguez-Iino (2021) and Schubert (2020). Unlike these papers, the overlapping generations structure of my model allows the CCPs to be computed relatively straightforwardly by backwards recursion. I separately estimate each city's housing supply elasticity using immigration shocks as an instrument for demand shocks. I then estimate the household utility parameters to maximize the (pseudo)-maximum likelihood function given data on household location choices.

The remainder of the paper is organized as follows. Section 1.3 provides some background to the changes to the US housing market. Section 1.4 presents the model. Section 1.5 describes the data. Section 1.6 reports on the parameterization. In section 1.7 I outline both the long-run and short-run counterfactual results. Section 1.8 concludes.

[3] Empirically Hornbeck and Moretti (2019) find that local productivity shocks induces inward migration and hence leads to rising rents and house prices, benefiting homeowners at the expense of renters.

1.3 Background

Prior to the 1970s land-use regulations were limited in scope, the right to build or demolish existing structures was by in large the right of the property owner. Since then, there has been a transformation of property rights, Glaeser et al. (2005). Owners often face significant obstacles when they wish to transform land from one use to another. Under the guise of historic preservation [4] existing structures cannot be demolished and environmental restrictions, such as urban growth boundaries, limit the supply of new land for residential structures. These changes were most significant in regions that subsequently had high-productivity growth, such as the San Francisco Bay Area and New York.

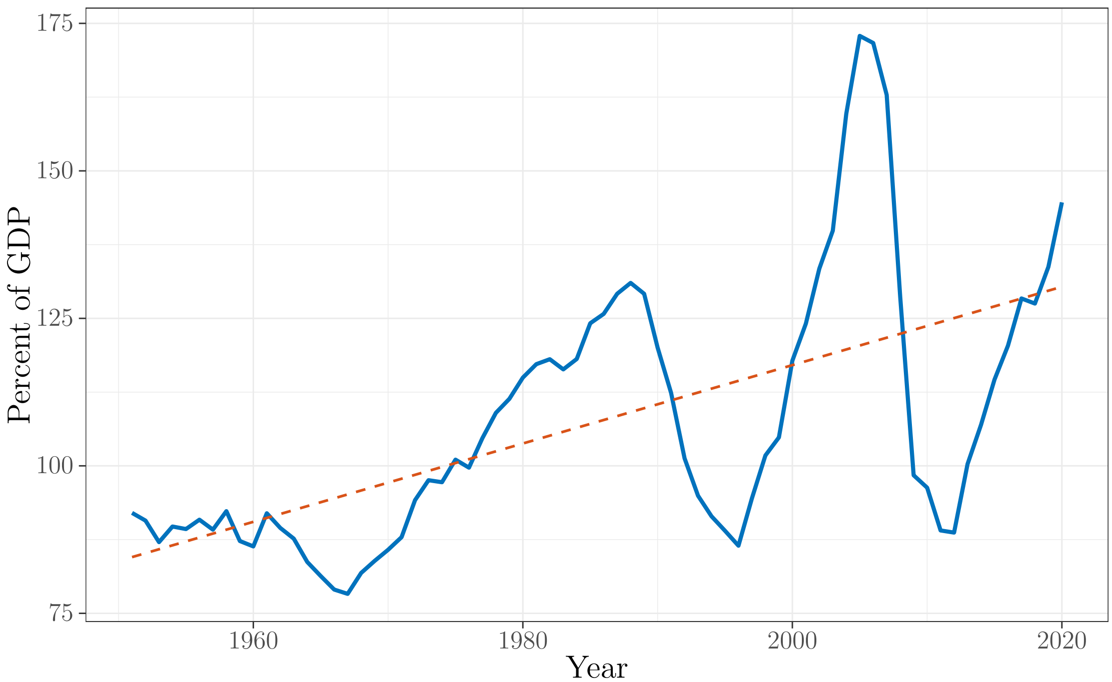
1.3.1 Housing has become expensive

The implication of these restrictions is that housing has become expensive relative to the cost of production. I show this fact in two ways. First, in Figure 1.1 I plot the difference between the market value and replacement value of real estate in the United States. While this series is volatile, the gap between the market and replacement cost of housing has increased since the 1980s.

One weakness of using land values is that time-varying discount rates may cause this series to change even in absence of any change in land-use regulations. When valuation ratios are high, the market value of housing will increase, even if there is no change in supply restrictions. Thus, a sequence of discount rate shocks could

[4]Land-use regulations are infamously strict in San Francisco. The city has effectively unlimited discretion in terms of approving new developments, which can lead to bizarre situations such as the preservation of laundromats from the 1980s. Kukura (2018)

1.1. Market value of residential real estate minus replacement cost.

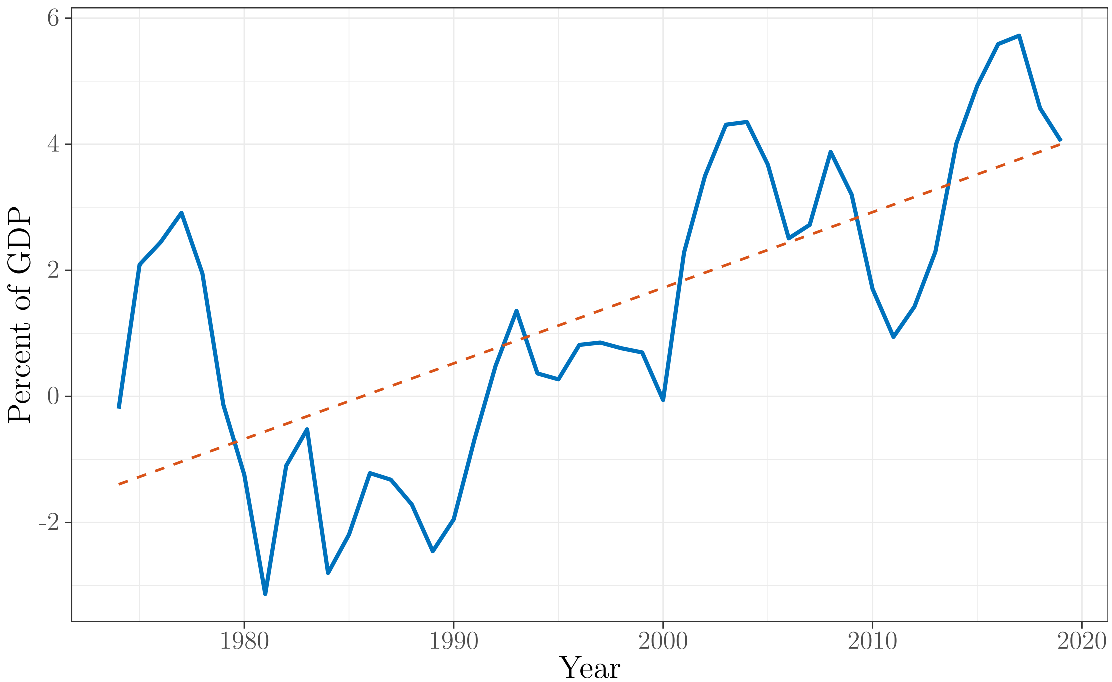


potentially cause the value of land to rise over time.

To rectify these concerns, I therefore also calculate the economic rents earned by the housing sector. Following Barkai (2020) I define the economic rents earned by the housing sector as the total value added by the housing sector (the value-added of housing services) minus payments to factors, which is principally the return on capital (structures), net taxes, depreciation and expected capital inflation. Since the value-added is derived from market rents, this measure is insensitive to the housing valuation ratios. I plot this series in Figure 1.2. This figure clearly shows that the economic rents of the housing sector relative to GDP has increased since the early 1990s. By the end of 2019, these economic profits were approximately 4% of GDP.

In both figures 1.1 and 1.2 we can see that housing has become expensive relative

1.2. Share of output that's attributed to land.



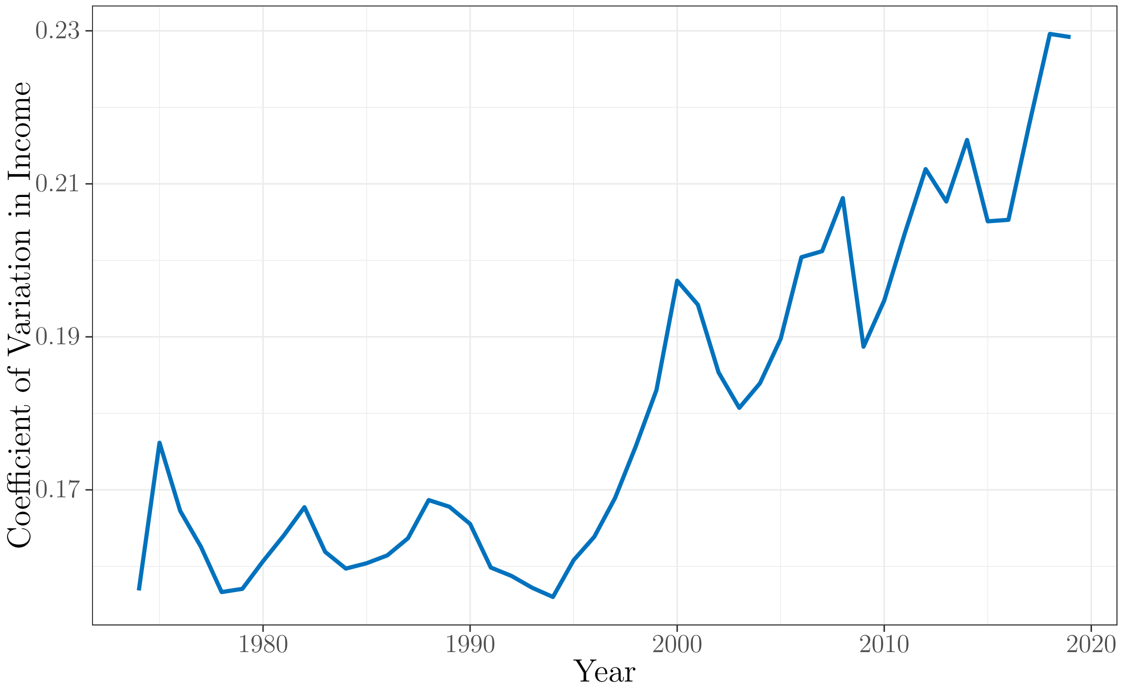
to the marginal cost of production[5].

1.3.2 Rising dispersion in incomes across cities

Secondly, in the data, we observe increased dispersion in incomes across cities. While for most of the post-war era incomes across regions converged, the opposite has been true in recent decades. In Figure 1.3 I plot the coefficient of variation in incomes across US cities from 1970s to present. The coefficient of variation is the standard deviation divided by the mean, and hence is a scale-free measure of dispersion. As we can see from the mid-1990s to the present day, the dispersion in

[5]Rognlie (2015) finds that the housing sector is responsible for much of the decline in labor share post-1980 in the US.

1.3. Coefficient of variation in mean household income across US cities.

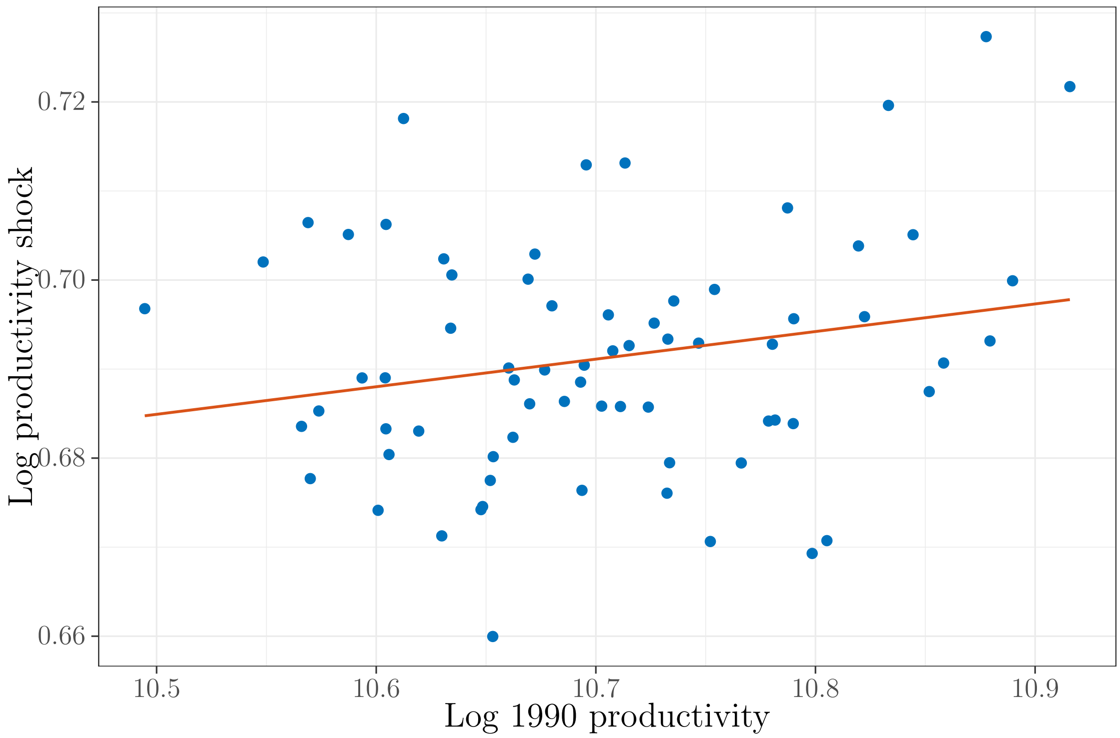


incomes between cities has increased by about 40%. Prior to the mid-1990s, however, income dispersion between regions was approximately constant.

Rising income dispersion between cities could be due to rising spatial sorting or cities with a higher initial level of productivity experiencing more rapid productivity growth [6]. In Figure 1.4 I plot the change in city-level productivity on the initial productivity level. As explained in detail in section 1.5, I calculate the initial productivity level as a city-level fixed effect from a regression of log individual wages on controls including industry and education of the worker. Productivity growth is

[6] Skill biased technological change or capital skill complementary, see Krusell et al. (2000), may also be responsible for rising income dispersion between locations, as cities differ in terms of their skill composition.

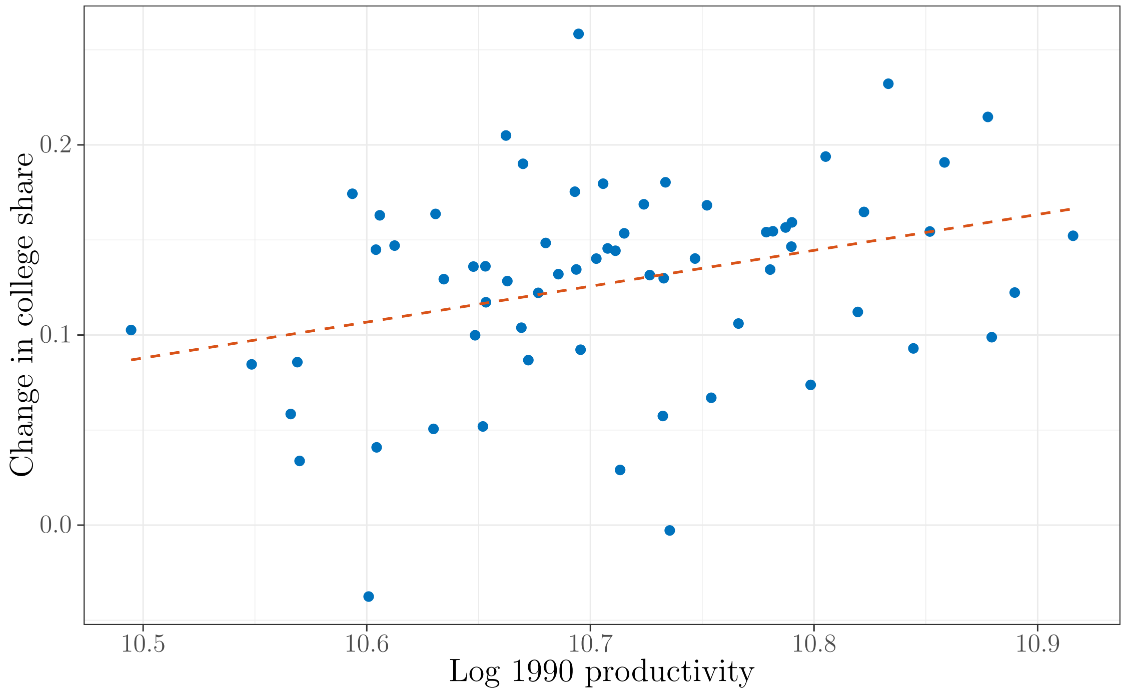
1.4. Change in city productivity on the initial productivity level.



defined as a Bartik shift-share instrument, where I fix the city industry wage bill shares at their 1990 level and use national changes in wages in each industry, excluding the city itself, and sum over all industries to calculate the city productivity shock. From the graph, we can see that there is a positive relationship between the city productivity level and its subsequent productivity growth, suggesting that rising income dispersion is in part due to uneven productivity growth.

Rising spatial sorting could also in part be responsible for the increased dispersion in incomes between cities. If there is an increased tendency for highly educated workers to live in the most productive cities, then the average wage in these cities

1.5. Change in the college share from 1990 to 2019 on the initial city productivity.

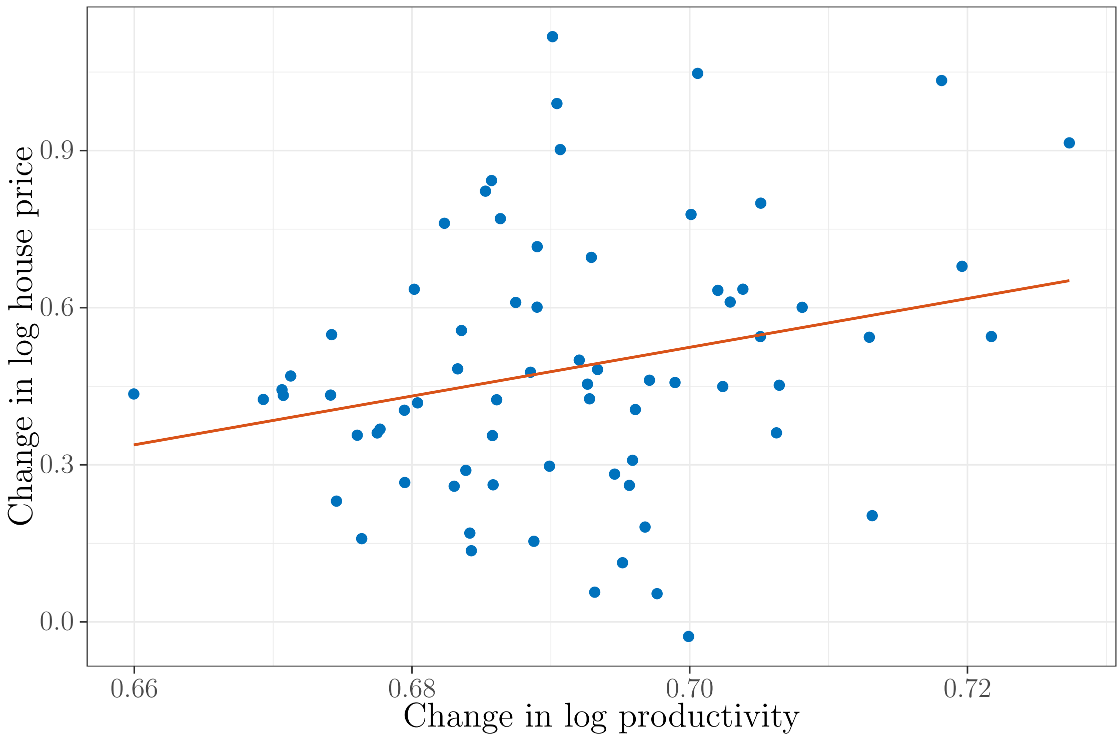


will rise relative to less productive locations. In Figure 1.5 I plot the change in the college-educated share of a city on its initial productivity level.

1.3.3 Positive correlation between income and productivity growth

In Figure 1.6 I show that that the change in log house prices and city productivity is positively correlated. Cities that experienced faster productivity growth also tended to have faster growth in house prices. I measure productivity growth as the Bartik change in wages for each industry and then scale this by the city wage bill. This Figure shows that the areas which had the greatest increase in wages became increasingly unaffordable. As a result, it has become increasingly difficult for younger

1.6. Change in log house price in a city against its change in log productivity.



and less educated households to move to these regions.

1.4 The Model

I consider an open economy, so that I take the interest rate as exogenous, with overlapping generations of households in a spatial equilibrium. The model captures the location and consumption savings decisions of richly heterogeneous households over their life-cycle. Households vary in terms of age, education, and wealth. Regions differ in local amenity supply, labor productivity, and housing costs. Housing costs are a function of local demand and the housing supply elasticity for the city.

1.4.1 Households.

Time is discrete, runs forever and is indexed by $t \in \{1, 2, \dots\}$. Households are either renters or homeowners. They are heterogeneous in terms of their age $a \in \{1, 2, \dots, A\}$, education $e \in \{1, 2, \dots, E\}$, risk-free asset holdings, $b \in \mathbb{R}$, current location $j \in \{1, 2, \dots, J\}$. The household's age and education determines their effective supply of labor, $f(a, e)$. Each period every household chooses a region j' to live in. When choosing a region, a household trades off moving costs, amenities, their idiosyncratic preference shock, and the regional productivity. A new generation of households are born each period and households die once they reach age A.

At the beginning of the period each household receives an i.i.d. vector of idiosyncratic preference shocks that are distributed according to a standard type one generalized extreme value distribution (G.E.V.). After receiving the preference shock the households choose a location j' to live in. Households receive flow utility from consumption, the local amenities, and their idiosyncratic preference shocks and incur moving costs when moving between regions. A household that previously lived in city j and chooses to live in city j' has flow utility

$$\underbrace{\log(c)}_{\text{Consumption}} + \overbrace{\omega\chi_{j'}}^{\text{Amenities}} - \underbrace{\tau_{jj'}}_{\text{Moving Costs}} + \overbrace{\theta\varepsilon_{ij'}}^{\text{Preference shock}} . \tag{1.1}$$

I assume that utility is separable in these terms and that household's receive logarithmic utility from consumption c. When the household lives in city j' their amenity consumption is determined by the local supply of amenities, $\chi_{j'}$ and their idiosyncratic preference for living in the location is determined by $\varepsilon_{ij'}$. The moving cost for a household previously in region j who moves to region j' is $\tau_{jj'}$[7]. The parameters ω and θ govern the relative importance of amenities and preference shocks to household utility. A larger θ (ω) increases the relative importance of preference shocks (amenities).

Households can choose to save in a risk-free asset or to borrow subject to a collateral constraint. I assume that all households start with no financial wealth and must end life with non-negative financial wealth. Households are renters for part of their life and homeowners otherwise. Specifically, I assume that households are renters prior to age $\overline{a}$ and then own housing after then. The Bellman equation of a household who owns a home, $a \in \{\overline{a}, \ldots, A\}$, is

$$V(a, b, e, j; S) = \max_{j', b'} \{\log(c) + \omega\chi_{j'} - \tau_{jj'} + \theta\varepsilon_{ij'} + \beta\mathbb{E}[V(a+1, b', e, j'; S')]\}, \quad (1.2)$$

where a is age, b is holdings of risk-free securities, e is education level, j is current location, and S is the aggregate state. The household chooses a new location j' and next periods bond holdings b' to maximize the right-hand side of (1.2). They discount the future at rate β and their expected value is computed over the support of preference shocks next period. Note that in the final period the expected continuation value equals zero, that is $\mathbb{E}[V(A+1, b', e, k; S')] = 0$. The households problem in

[7] I assume that there is no moving cost in the first period of life. This ensures that where the households are born doesn't affect the steady-state. If there were moving costs in the first period then the distribution of where households are born would matter for the equilibrium.

(1.2) is maximized subject to following constraints

$$c + p_{j'} + \frac{b'}{1+r} = (1-\delta)p_j + w_{j'} f(a,e) + b, \tag{1.3}$$

$$b' \geq -\psi p_{j'}, \tag{1.4}$$

$$S' = G(S),\ \varepsilon_{ij} \overset{iid}{\sim} G.E.V.(1). \tag{1.5}$$

Equation (1.3) is the home owner's budget constraint. The household uses these resources to purchase risk-free securities, b', with interest rate r, amount c of the consumption good, and housing in their current location j'. The price of an owner occupied home in their current location is $p_{j'}$. The right-hand side of the budget constraint consists of the depreciated value of their home in the previous period [8]. The depreciation rate of housing is δ per period. Additionally, households resources include the wage, per effective unit of labor, of region j' multiplied by their effective supply of labor, as well as their current holdings of risk-free securities.

All debt must be collateralized and there exists a down-payment requirement of $1-\psi$ when purchasing a home. This gives rise to the borrowing constraint in equation (1.4) [9]. The aggregate state is denoted by S and evolves according to the function G. I defer further discussion of the aggregate state until I define the equilibrium.

For renters, that is households with $a \in \{1, 2, \ldots, \overline{a}-1\}$, their maximization problem is the same as in (1.2). However, the resource constraint, (1.3), and borrowing

[8] If the household was a renter in the previous period then this term is not in their budget constraint.

[9] Since all households end life with no wealth I allow them to consume their down-payment in their final period of life. Thus, I drop (1.4) and the budget constraint in the final period is $c + p_{j'} = (1-\delta)p_j + w_{j'} f(a,e) + b$

constraint, (1.4), are now given by the following two equations

$$\frac{b'}{1+r} + c + p_{rj'} = w_{j'} f(a,e) + b, \tag{1.6}$$

$$b' \geq 0, \tag{1.7}$$

where $p_{rj'}$ is the price of rental housing in region j'.

Note that the expectation in (1.2) is taken over the possible realizations of idiosyncratic taste shocks next period. It is helpful to write the conditional value of living in location j', *excluding* the realized taste shocks as $W(a,b,e,j,j';S)$ and then continuing optimally after then. That is

$$W(a,b,e,j,j';S) = \max_{b'}\{\log(c) + \omega\chi_{j'} - \tau_{jj'} + \beta\mathbb{E}[V(a+1,b',e,j';S')]\}. \tag{1.8}$$

Which implies that we can write equation (1.2) of the households problem as

$$V(a,b,e,j;S) = \max_{j'}\{W(a,b,e,j,j';S) + \theta\varepsilon_{ij'}\} \tag{1.9}$$

Given the assumed distribution of the preference shocks we can write the probability that a household of type (a,b,e,j) chooses location j' as

$$\pi(j'|a,b,e,j;S) = \frac{\exp(W(a,b,e,j,j';S))^{\frac{1}{\theta}}}{\sum_k \exp(W(a,b,e,j,k;S'))^{\frac{1}{\theta}}}. \tag{1.10}$$

Furthermore, by standard results, Rust (1987), Hotz and Miller (1993), and the assumed distribution of the preference shock implies that the expected value of the household prior to the realization of the taste shocks can be expressed as

$$\mathbb{E}[V(a,b,e,j;S)] = \max_{b'}\{u(c,j,k;S) + \theta\gamma - \theta\log(\pi(k|a,b,e,j;S)) + \beta\mathbb{E}[V(a+1,b',e,k;S')]\}, \tag{1.11}$$

$$u(c,j,k;S) = \log(c) + \omega\chi_k - \tau_{jk}, \tag{1.12}$$

where k is an arbitrary location and γ the Euler-Mascheroni constant. The importance of equations (1.8) to (1.11) are in aiding the computation of the household's conditional choice probabilities, $\pi(j'|a,b,e,j;S)$. To compute the probability that a household chooses a location we insert the solution of (1.8) into (1.10). However, to compute (1.8) we need next periods expected value, which is taken over realizations of the households preference shocks. Conveniently, the assumption that households preference shocks are distributed G.E.V. (Arcidiacono and Ellickson (2011)) implies that the expected value has the tractable form in (1.11), which can be solved by backwards recursion and using the fact that $\mathbb{E}[V(A+1,b',e,k;S')] = 0$.

1.4.2 Housing Sector

Housing is produced in each city using the consumption good and the local productivity of producing housing depends on the city specific supply elasticity. Each construction firm is competitive, however at the city-level there are decreasing returns to scale relative to the aggregate housing stock: as the density of the city increases the productivity of the construction sector falls. This captures the fact that regulatory constraints on new construction tend to limit density or preserve existing low density areas. Motivated by this I therefore assume that new housing is produced using the following production function

As argued in Glaeser and Gyourko (2018) construction costs are approximately linear in the floor area of a structure beyond a certain threshold. Thus, the gap between replacement costs and market values does not reflect technological constraints such as decreasing returns to scale in the construction sector, but rather regulatory constraints on new housing.

$$Y_{hjt} = A_h \left(\frac{\overline{H_j}}{H_{jt}} \right)^{\alpha_j} x_c. \tag{1.13}$$

Here x_c is the quantity of consumption good used, and is the only input in constructing new housing. The aggregate housing stock in the city is given by H_{jt} and $\overline{H_j}$ is the land area in the city that's suitable for construction. Thus, the density of the city is $\frac{H_{jt}}{\overline{H_j}}$. Each city has it's own city specific supply elasticity, α_j, [10] this determines how responsive city-level prices are to changes in local population. This supply elasticity is implicitly determined by the local land-use regulations that make it costly to construct new housing in more regulated areas. As α_j are positive as a city's density increases, the productivity of the construction sector declines. The law of motion for the housing stock is given by

$$H_{jt} = (1 - \delta) H_{jt-1} + Y_{hjt}, \tag{1.14}$$

where δ is the economy wide depreciation rate for housing and Y_{jt} is the quantity of

[10] Throughout my paper I assume a single housing supply elasticity for each urban area that is constant over time. My framework could easily be extended to incorporate time varying housing supply elasticities. This could capture the changes in land-use regulation as documented in Gyourko et al. (2021).

new housing produced in city j. We can write the construction firm's problem as

$$\max_{x_c} p_{jt} A_h \left(\frac{\overline{H_j}}{H_{jt}} \right)^{\alpha_j} x_c - x_c, \tag{1.15}$$

where we note that the price of the consumption good is normalized to one.

1.4.3 Rental Sector

There are no frictions in transforming owner occupied housing into rental housing. The fraction of housing in a city that is rental housing is thus an equilibrium object. To simplify matters I assume that the rental housing stock is owned by risk-neutral foreigners with deep pockets. This avoids the need for an additional state variable in the households problem as well as simplifying the market clearing conditions [11]. I assume that these foreigners discount the future using the same risk-free rate r that the household receives on their risk-free bond holdings. In each city j their problem is

$$\max_{q_{jt} \in \mathbb{R}_0^+} \left(p_{rjt} + \left(\frac{1-\delta}{1+r} \right) p_{jt+1} - p_{jt} \right) q_{jt}. \tag{1.16}$$

In equation (1.16) q_{jt} is the quantity of rental housing owned the foreign entity in city j at time t. The rental firm solves this problem for each city j.

1.4.4 Consumption Good Sector.

All workers in city j work in the consumption good sector of city j. The consumption goods sector in each region has a region-specific productivity A_{jt} that is

[11] In steady-state these foreign entities earn no economic profits and hence the importance of this assumption to inequality is largely along transitions between steady-states.

common to all firms in the region. The region-specific productivity is non-stochastic for simplicity. A continuum of competitive firms use local labor to produce tradable consumption goods. The representative competitive firm f produces output according to

$$A_{fjt}N^e_{fjt}, \tag{1.17}$$

where N^e_{fjt} is the effective units of labor in region j at time t that are employed by firm f. The total measure of effective units of labor in region j is simply $\sum_{e,a}\mu(a,e,j;S)f(a,e)$, the total measure of workers of age a and education e, $\mu(a,e,j;S)$, times their effective supply of labor, $f(a,e)$ [12].

1.4.5 Equilibrium

I solve for a steady-state equilibrium, that is, an equilibrium where all prices value functions, policy functions, and distribution are constant over time.

The wealth of the household will depend on their age and education as well as their history of location choices. Let this endogenous distribution of household bond wealth conditional on age, education, location prior to moving, be denoted by $\Psi(db|a,e,j)$. I label the measure of households of type a,e in region j as $\mu(a,e,j)$.

A steady-state equilibrium is defined as a set of prices, $p_j, j \in \{1,\ldots,J\}$, for housing in each region j and a set of prices for rental housing in each region $p_{rj}, j \in \{1,\ldots,J\}$, a value function, $V(a,b,e,j;S)$ for the household, a policy function for the household $b'(a,b,e,j,j';S)$, conditional choice probabilities for the household, $\pi(j'|a,b,e,j;S)$, a household wealth distribution, $\Psi(db|a,e,j)$, and $\mu(a,e,j)$ mea-

[12] In this study I assume that there are no local agglomeration economies. Introducing agglomeration economies into the model can lead to multiple steady-states.

sures of households across locations, such that,

1. Given the set of prices all agents choose optimally. That is the households solve their problem, (1.2) to (1.7), the construction firm solves the problem given by (1.15) and the foreign owners of the rental housing solve their problem, in equation (1.16). For the household their value function $V(a, b, e, j; S)$ and policy functions $\pi(j'|a, b, e, j; S)$ and $b'(a, b, e, j, j'; S)$, jointly solve their Belman equation.

2. Given the solutions to the agent's problems all markets clear.

3. The measure $\mu(a, e, j)$ is consistent with the policy functions $\pi(j'|a, b, e, j; S)$ and the wealth distribution $\Psi(db|a, e, j)$.

4. $\Psi(db|a, e, j)$ is consistent with the policy function $b'(a, b, e, j, j')$ and the probabilities $\pi(j'|a, b, e, j; S)$.

Since this is an open economy so there is no market clearing condition for the bond market.

1.4.6 Solution

Given the complexity of the model it is of no surprise that there is no closed form solution and so I resort to numerical methods. To simplify this, I first derive some steady-state equilibrium relationships for the housing rents and the price of housing.

Since the rental housing stock is owned by a competitive foreigner with deep pockets a no arbitrage condition between rental and owner-occupied housing must hold. That is the price of a home must also equal the present discounted value of future rents. If we let the price of rental housing in city j be p_{rjt} then it must be that

$$p_{rjt} = p_{jt} - \left(\frac{1-\delta}{1+r}\right) p_{jt+1}. \tag{1.18}$$

Equation (1.18) comes directly from solving the rental firm's problem. Using this we can easily solve for the price of rental housing in terms of the owner occupied housing. Note that in steady-state since all prices are constant this implies that $p_{rj} = \frac{r+\delta}{1+r} p_j$.

Furthermore, the competitive nature of the construction sector means we can solve for the representative construction firm's problem to solve for house prices in each region in terms of the measure of households in the region. It is thus easy to show that the equilibrium house prices in each region is given by

$$p_j = \frac{1}{A_h} \left(\frac{H_j}{\overline{H}_j}\right)^{\alpha_j}, \tag{1.19}$$

where $H_j = \sum_{a,e} \mu(a, e, j)$, is the measure of households in region j at time t. Again, in steady-state this measure must be constant. In equation (1.19) we can see that there is an increasing relationship between the price of housing in a region it's population density. As density increases regulations that limit the supply become more important and hence the price of housing will begin to rise.

Given a set of prices one can compute the conditional value function for a household living in location j', *excluding* the realized taste shocks (i.e. $W(a, b, e, j, j'; S)$, and call this the conditional value function), by backwards recursion. Beginning in the final period of life we know that $\mathbb{E}[V(A+1, b', e, k; S')] = 0$, so the conditional value function can be computed easily. We can then use the conditional value function to calculate conditional choice probabilities and use these probabilities to evaluate the expected value function in the next to last period, $\mathbb{E}[V(A, b', e, k; S')]$, as in (1.11). This allows us to solve the households conditional value function in the second to last period. I continue this process until we reach the period in which the first period of life.

Specifically, I discretize the bond grid and interpolate over this grid. The remaining state variables a, e, j, j' are discrete and therefore are exactly represented on the grid. Note however, that even conditional on a, e, j we have a non-degenerate wealth distribution. This is because the entire history of locations that were visited matters for total earnings and spending on housing. Thus, the true endogenous wealth distribution of the model is a high dimensional object and cannot be represented exactly [13].

Following Krusell and Smith (1998) I approximate the true wealth distribution using a lower dimensional object. While it is tempting to use the first few moments of the unconditional wealth distribution to approximate the true wealth distribution this does not appear to accurately represent the true distribution well. This is likely due to the highly heterogeneous nature of the households. Wealth depends on age, education, location as well the history of locations the household visited.

Motivated by this I use the mean of wealth conditional on age, education and current location as my low dimensional approximation of the true wealth distribution. This removes the location history dependence from the wealth distribution. Simply put I assume that $b(a, b, e, j)$ is a constant in b. I solve for prices where households

[13] In my quantitative exercise the true endogenous wealth distribution will have approximately 3.23×10^{15} points in steady-state, and hence is not computable.

treat the mean level of wealth conditional on a, e, j as the true distribution[14].

I am thus left with a fixed-point problem where we need to solve simultaneously for equilibrium prices and equilibrium measures in each region. Computationally, I start with an initial price guess for the price of housing in each city. Then using this price guess, I solve the households problem and hence the implied measure of households in each city. The housing supply, in (1.19), maps these measures to implied supply prices. I then update the price guess using these supply prices. I repeat this until I converge to a predetermined tolerance. A more complete description of the solution method is in Appendix (1.A).

1.4.7 Model Discussion

The model in this paper differs from Diamond (2016) or Schubert (2020) in several key ways, one of which I highlight here. In my paper, households can be either homeowners or renters in my model and have a dynamic income-consumption problem. They can smooth their consumption and housing expenditures over their lifetimes using the risk-free asset. In my model households begin adult life as renters with little wealth, then save for a down payment, and become homeowners in middle age. This creates a life-cycle pattern of housing wealth and differences in the portfolio composition across households. These differences in housing wealth is one potential reason the impact of changes to land-use regulations might be heterogeneous across households.

In equilibrium housing demand is increasing in the region's productivity and

[14]When I simulate the model using a large number of households using the steady-state prices I find that the measure of households that are in each city to be very close to the measures I found using the approximate wealth distribution.

amenity supply. A region with high-productivity or amenities is a more desirable location to live and so an all-else equal rise in either local productivity or amenity supply will increase housing demand in the region. Therefore, except in the case that the local housing supply is infinitely elastic, the equilibrium house price will be increasing in local productivity and local amenity supply. However, differences in local housing supply elasticities and imperfect correlation between amenity supply and productivity will mean that there is not a perfect rank correlation between city productivity or amenity supply and house prices.

The model generates assortative matching between a worker's education and the city's productivity, in the sense that highly educated workers are more likely to locate in high-cost high-productivity cities. I show this in a simplified version of the model in Appendix 1.C. Intuitively, since workers demand only one home, as a worker's income increases the expenditure share on housing declines and so they are better able to afford the high-cost high-productivity locations [15]. As equilibrium house prices are typically increasing in the regions productivity the model generates sorting along education and city productivity lines.

My model also generates a similar pattern of sorting along education and amenity lines. Since more educated households have a smaller marginal utility from consumption, they will tend to sort into expensive locations with a high amenity supply. Again, all else equal housing costs will be increasing in the local amenity supply and so the wealthy will tend to locate in high amenity locations. However, the presence of idiosyncratic preference shocks means that the model does not generate perfect sorting along education and city productivity lines, as we see in

[15] In the data I find that a 10% increase in household income increases expenditure on housing by about 2% when a city-level fixed effect is included.

Van Nieuwerburgh and Weill (2010). Some highly educated households will select less productive areas as their idiosyncratic preferences for living in a location are sufficiently large to offset the reduced consumption or amenity utility.

1.5 Data

This paper uses the Core-based Statistical Area (CBSA) from the Office of Management and Budget as its definition of a city. The paper's main data source is the US Census 5% microdata samples from IPUMS (Ruggles et al. (2021)). My sample consists of workers aged 25-66 living in US metropolitan areas. I aggregate counties into CBSAs using the NBER crosswalk where appropriate. I estimate the model taking 1990 as a steady-state. Summary statistics for the Census sample data are in table 1.1

1.1. Summary statistics of the household level data.

Statistic	N	Mean	St. dev.	Median	Pctl(25)	Pctl(75)
Annual Wage	2,262,973	25,757.790	25,371.510	21,000	10,381	34,000
Hours worked	2,262,973	39.670	13.197	40	40	45
Age	2,262,973	40.284	10.484	39	32	48
Migrated	1,798,742	0.117	0.321	0	0	0

Wages: From the 5% microdata samples I obtain wages for individuals. This gives wages conditional on location, age, education, worker-industry, and hours worked.

Migration: The 5% microsample includes the location of the worker five years ago. International migration rates and annual total city population comes from the Census Bureau.

House prices: A quarterly index panel of house prices for the top 100 cities is available for US cities beginning in 1990 from the Federal Housing Finance Authority (FHA). I use this data as it is based in actual transactions rather than appraised housing values. I use the 1990 5% sample to get the initial level of house prices in each city. Using the FHA data I obtain the log annual house price growth for each city by taking the log annual difference in the house price index over time [16]. I combine this with the land-use regulation index from Gyourko et al. (2008) and the fraction of land that's available for development from Saiz (2010).

City land area: I obtain data on the city size from the US Census Bureau. I define the $\overline{H}_j$ of the city to be the metropolitan land area in square miles and I normalize this measure so that the total land area sums to one. These city-level summary statistics are in table 1.2.

1.2. City-level summary statistics.

Statistic	N	Mean	St. dev.	Pctl(25)	Median	Pctl(75)
Δ log(Population)	2,130	0.01	0.01	0.005	0.01	0.02
Inter. migration per 1000	2,130	2.85	2.39	1.22	2.20	3.59
Δ log(House prices)	1,740	0.04	0.06	0.02	0.04	0.07
City Amenities	2,400	−0.61	1.60	−1.50	−0.62	0.42
Real wage shock	2,059	0.02	0.01	0.02	0.02	0.03
Population 1990	71	2.11	2.67	0.70	1.31	2.46
Productivity 1990	71	44,681.83	4,190.94	41,530.12	44,121.87	47,688.69
City land area	71	711.26	632.09	313.18	523.03	779.86
Percent of land unavailable	71	25.44	21.07	9.04	19.05	38.22
Land regulation index	71	0.11	0.70	−0.38	0.05	0.54

Amenities: Local amenities supply consists of per capita supply of services establishments, from the Quarterly Census of Employment and Wages. As described

[16] All prices are deflated to their 1990 values using the personal consumption expenditures index from the BLS.

in section 1.6 I combine this with per capita violent crime from the FBI to create the local amenity supply of the city using a principal components analysis. Summary statistics are in table 1.3.

1.3. Summary statistics of the local amenities.

Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
Art-recreation	2,400	0.347	0.109	0.284	0.335	0.398
Drinking places	2,400	0.442	0.110	0.365	0.429	0.504
Restaurants	2,400	0.253	0.081	0.196	0.245	0.304
Grocery stores	2,400	0.153	0.094	0.081	0.136	0.199
Cinemas	2,400	0.013	0.006	0.010	0.013	0.016
Clothing stores	2,400	1.426	0.267	1.239	1.432	1.606
Violent crime	2,400	1.744	1.100	1.050	1.586	2.316

Local Productivity Growth: Using the Quarterly Census of Employment and Wages I create Bartik productivity shocks to calculate the changes in productivity in each region. Specifically, I use the 1990 county-level three-digit NAIC industry wage-bill shares [17] and aggregate these up to the CBSA level[18]. In constructing the national productivity change for industry i for city j I exclude the city j's contribution to that national change. Thus, the productivity shock for industry i in city j is the product of the 1990 wage bill share times the all but j change in real hourly wage in industry i. The total productivity shock for city j is then the sum of wage bill for industry i times the all but j change in real hourly wage for industry i. Specifically,

[17] In calculating the productivity shock I hold the wage bill shares fixed at their 1990 level, as the industry composition is likely endogenous to the skill makeup of the area. So endogenous changes in the skill make-up the region could affect the measurement of productivity.

[18] I aggregate up from the county level to the CBSA level as when a lower unit is censored for anonymity reasons all higher units associated with it be censored. Thus, aggregating from lowest level unit (counties) leads to the least amount of censoring.

the wage bill shares for industry i in city j is given by

$$WB_{ij} = \frac{\text{Hourly wage } 1990_{ij} \times \text{Hours } 1990_{ij}}{\sum_k \text{Hourly wage } 1990_{kj} \times \text{Hours } 1990_{kj}}. \tag{1.20}$$

Given these wage bill shares the productivity shock for city j in year $t+1$ is given by

$$\Delta \log(\tilde{w}_{jt+1}) = \log(\sum_i WB_{ij} w_{i-jt+1}) - \log(\sum_i WB_{ij} w_{i-jt}), \tag{1.21}$$

where w_{i-jt} is the all but j hourly wage for industry i in year t. Table 1.2 contains summary statistics for the city-level data used in the paper.

1.6 Parameterization

For the quantitative exercises I parameterize the model so that one period is five years as this maps easily into the Census data on intercity migration. I assume that education takes five values, less-than high-school, high-school, some-college, bachelors, and more than bachelors. I develop an estimate for local amenity supply, estimate the housing supply elasticities and preference parameters of the workers. The remaining parameters of the model are either calibrated from the literature or estimated directly.

I parameterize the model so that one period represents five years, as intercity migration data conditional on education and age is at the five-year frequency. Life begins at age 25 and ends at age 65. I reduce the dimension of amenities to a single variable, calibrate and estimate the model's parameters.

1.6.1 Amenities

To calculate amenities, I use per person establishment counts from the Quarterly Census of Employment and Wages. The following amenity producing industries are counted Arts, Entertainment, and Recreation NAICS 71, Drinking places, NAICS 72241, Restaurants and Other Eating Places, NAICS 72251, Grocery Stores, NAICS 4451, Motion Picture and Video Exhibition, NAICS 51213, and Clothing and Clothing Accessories Stores NAICS 448. Per capita violent crime data comes from the FBI.

Since amenities are assumed to have only one dimension in the model and that many of these amenities are correlated with each other, I follow Diamond (2016) and perform a principal components analysis (PCA) on the amenities data, pooling all the data from all time periods. I scale the data so that all variables have a mean of zero and a standard deviation of one. The loadings on the first principal component are in table 1.4. The value amenities in city j at date t is the sum of the product of the loading on each variable on the first principal component times the value of the variable in the city at time t.

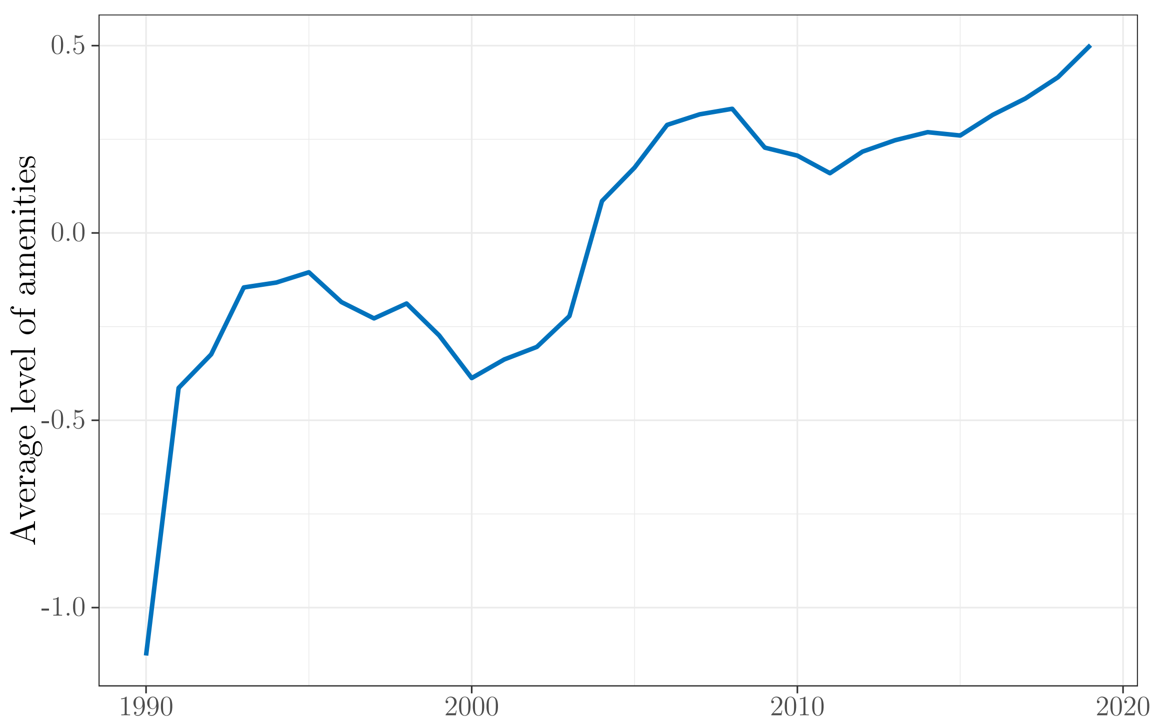
1.4. PCA loadings used in the calculation of amenities.

Variable	Loading
Art-recreation	0.414
Drinking places	0.459
Restaurants	0.299
Grocery stores	0.214
Cinemas	0.383
Clothing stores	0.573
Violent crime	−0.092

As we can see in table 1.4 the PCA has a positive loadings on the per-capita number of amenity establishments in the city. Thus, as one would expect a higher number of service establishments in a city means the amenity supply of the city is larger. We see that there is a negative loading on violent crime, indicating that cities with more violent crime have lower amenities. Again, this is in line with economic intuition as crime is a negative amenity.

In Figure 1.7 I plot the log level of amenities over time. As we can see the average level of amenities has increased over time, reflecting increased per person services supply and falling violent crime.

1.7. Mean level of the log level of amenities across cities.



1.6.2 Calibrated Parameters

The assumption that there is no default and that debt must be fully collateralized implies in steady-state that the borrowing constraint must satisfy $\psi \leq 1 - \delta$. A borrowing limit greater than this would mean that the household's debt at the beginning of the next period would be greater than the value of their collateral. In addition assume that the lender doesn't recover the full value of the home. Evidence from Campbell et al. (2011) suggests that homes sold as part of a forced sale are sold below their market value. I therefore follow their paper and assume a 6.5% loss when the borrower defaults, which is their estimate for the loss in value when a home is sold around a bankruptcy by a single seller.

As is well known in the literature, e.g. Arcidiacono and Ellickson (2011), it is difficult to estimate the discount factor with dynamic discrete choice models. I therefore assume that $\beta = \frac{1}{1+r}$ so that the discount rate of households equals that of the foreign investors. I set the real-interest rate equal to the post-war risk-free rate of 1.3% annual, Jordá et al. (2019), and hence the discount factor equals $(0.987)^5 = 0.9366$. I set the annual depreciation rate on housing to equal 0.77%, Davis and Heathcote (2007), which implies that that per period depreciation rate equals 3.9%.

A final model parameter that is calibrated is A_h, the productivity at which consumption is converted to housing. In calibrating this parameter, I target the average house-price across cities in 1990 and use the estimated housing supply elasticities that are derived later. After estimating α_j for each city I use the IPUMS data I calculate the measure of households in each region and then using the estimated elasticities and city sizes I can calculate the predicted house price in each region. Using the equilibrium condition for house prices I set A_h to match the population weighted mean of house prices across cities. That is $A_h = \sum_j \frac{1}{p_j}\left(\frac{H_j}{\overline{H_j}}\right)^{\alpha_j} \frac{Pop_j}{\sum_k Pop_k}$. All calibrated parameters are in table 1.5. [19].

1.5. Calibrated model parameters.

Parameter	Description	Value	Source
$\beta = \frac{1}{1+r}$	Discount rate	0.928	Jordá et al., 2019
ψ	Borrowing limit	0.899	No-default assumption.
δ	Housing depreciation rate	0.039	Davis and Heathcote (2007)
A_h	Construction sector productivity	0.4851	1990 House-prices

1.6.3 Estimating City Productivity and Effective Labor Supply

Since workers sort into cities, it is difficult to estimate the initial level of productivity, as more productive regions will tend to have workers who are more educated on average. This positive assortative matching between the worker's effective supply of labor and the city's productivity means that comparing average hourly wages between regions will tend to overstate the productivity differences between them.

From the labor consumption goods equilibrium we the wage of a worker of age a and education e in city j is given by $f(a, e)$

$$Wage_{jae} = A_j f(a, e). \tag{1.22}$$

Taking logs of 1.22 and assuming that age and education have separable effects

[19] It is possible to allow A_h to vary between cities reflecting underlying construction productivity differences, e.g. due location specific technologies, such as the need for buildings to be resistant to earthquakes in California. For this paper however I assume that the construction sector productivity is constant across regions.

on wages [20], that is $f(a,e) = g(a)s(e)$. Then I can write the log wage of worker l with education e working hours h per week living in city j as

$$\log(Wage_{ljeiah}) = \delta_j + \delta_e + \delta_a + \delta_i + \delta_h + \varepsilon_{ljeiah}. \tag{1.23}$$

where I include an additional fixed effect, δ_i, for the worker's industry. The estimates of δ_j are the city's 1990 productivity level. The estimates of δ_e are used to calculate the effect of education on wages and represent our estimate of $s(e)$. The δ_a will be used to capture the effect of age on earnings, that is our estimate of $g(e)$. The remaining terms are controls that don't map explicitly to model objects. The results of this regression are in table 1.6.

1.6.4 Estimating Elasticities

A key parameter that needs to be estimated is the housing supply elasticity. The housing supply elasticity determines the responsiveness of local house prices to a change in the population (or equivalently the housing stock) of the city. I allow elasticities to vary across cities reflecting how the supply elasticity varies across regions but assume that this elasticity are a function of local land-use regulation.

From the construction firm's first order condition I can solve for equilibrium price of housing in each city. The inverse housing supply relationship is:

$$p_{jt} = \frac{1}{A_h}\left(\frac{H_{jt}}{\overline{\overline{H}}_j}\right)^{\alpha_j}, \tag{1.24}$$

[20] Including an interaction term between the workers age and education does little to affect the measurement of initial city productivity. Results available upon request.

1.6. Initial city productivity regression.

Dependent Variable:	$\log(Wage)$
Model:	(1)
Variables	
DNF High-School	-0.6283***
	(0.0179)
High-School	-0.4102***
	(0.0152)
Graduate degree	0.2049***
	(0.0271)
Some College	-0.2618***
	(0.0097)
Fixed-effects	
City	Yes
Age	Yes
Hours	Yes
Industry	Yes
Fit statistics	
Observations	2,797,254
R^2	0.43675
Within R^2	0.08411

Two-way (City & Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

which determines the price of housing in each city. Note that by taking the derivative of the log of 1.24 that we can see that the housing price elasticity for city j is equal to

$$\frac{1}{\alpha_j}. \tag{1.25}$$

To estimate the value of α_j in each city I take log differences of (1.24) and obtain

$$\Delta \log(p_{jt}) = \alpha_j \Delta \log(H_{jt}) + \Delta \log(A_{ht}) + \Delta u_{jt}. \tag{1.26}$$

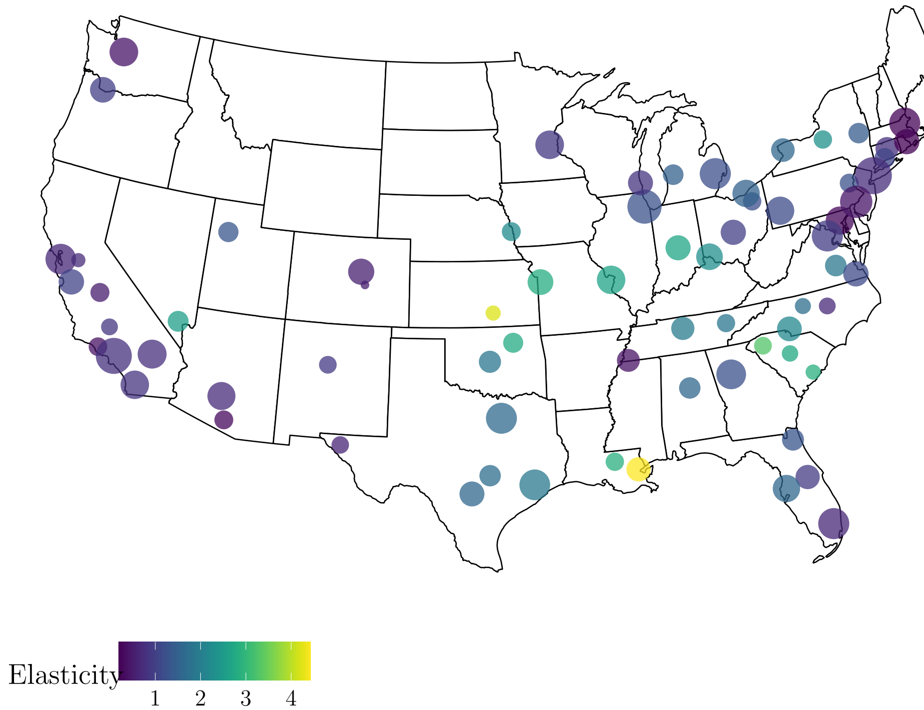
In (1.26) I allow the construction sector productivity to vary across regions and over time, as Δu_{jt} may be non-zero. Regressing $\Delta \log(p_{jt})$ on the $\Delta \log(H_{jt})$ may lead to inconsistent estimates of α_j as the error term, Δu_{jt}, could be correlated with $\Delta \log(H_{jt})$. This is because construction sector supply-side shocks would change equilibrium quantities and prices. For example, an improvement in local construction sector productivity would likely reduce the equilibrium house price and hence increase the regional population. Hence, using OLS to estimate (1.26) would lead to inconsistent estimates for α_j.

To alleviate these endogeneity issues, I propose using international migration as an instrument for local housing demand shocks. The literature has argued that international migration takes place along established networks reflecting the cultural nature of migration Saiz (2010). This implies that international migration patterns are unrelated to local construction sector productivity and hence that they are a valid instrument for H_{jt} [21].

As argued by Glaeser and Gyourko (2018), Glaeser et al. (2005), and Quigley and Rosenthal (2005) beyond a certain threshold construction costs are approximately linear in floor area. Constructing a ten story building costs approximately twice as much as constructing a five story structure and so without any land-use restrictions

[21] Diamond (2016) and Schubert (2020) instrument for housing demand using shift-share productivity shocks similar to what I constructed in section 1.5. However, recently Goldsmith-Pinkham et al. (2020) raises questions as to the validity of shift-share instruments as one would need to assume that the initial industry shares are exogenous.

1.8. Housing supply elasticity by city.



there should be no relationship between density and house prices. As these papers argue, by restricting the supply of new housing land-use regulations increases the price of housing in a region. Motivated by this, and similar to Diamond (2016) I thus assume that local housing supply elasticity depends on the level of land-use regulations within a city. I therefore that the local housing supply elasticity has the following functional form[22]:

$$\alpha_j = \alpha \exp(Regs_j), \tag{1.27}$$

[22]With a linear functional form for α_j I find negative elasticities for a small number of regions.

where $Regs_j$ measures the stringency of land-use regulations of city j, from Gyourko et al. (2008). The value of α_j used is the predicted value from our regression. In addition to the main variable of interest, the interaction of exponential of regulations and log population changes, I include the log of the total population and the log of the land that is not available for development as controls in the regression. The results of this regression are in table 1.7. Figure 1.8 plots the geography of the estimated supply elasticities.

1.7. Regression used to estimate the city-level housing supply elasticities.

Dependent Variable:	$\Delta \log(price)$
Model:	(1)
Variables	
(Intercept)	-0.0460*
	(0.0272)
exp(Regulation) $\times$ $\Delta Population$	0.7823***
	(0.2924)
$\log(Unavailableland)$	0.0011
	(0.0013)
$\log(Population)$	0.0033*
	(0.0019)
Fit statistics	
Observations	1,951
R^2	0.01761
Adjusted R^2	0.01610

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

1.6.5 Estimating Household Utility Parameters

I first estimate city productivity levels and growth, the city housing supply elasticity. Then after determining the level of amenities in each city and calibrating the parameters in table 1.5 I estimate the household utility parameters.

The household's utility parameters, don't map easily to the data. I therefore structurally estimate these parameters by maximizing the likelihood of the model, Rust (1987), Hotz and Miller (1993), Aguirregabiria and Mira (2002), Arcidiacono and Miller (2011). In my case given a set of model parameters I can calculate the CCP of the household choosing location j' given their state. These CCPs form the the (pseudo) likelihood function of the model and model parameters are selected to maximize this likelihood function.

In the structural estimation I assume that the moving cost has the following functional form

$$\tau_{jj'} = \begin{cases} 0 \text{ if } j = j' \\ \tau \text{ otherwise.} \end{cases} \tag{1.28}$$

1.6.5.1 Comparative Statics:

Using a likelihood method to estimate the model means all moments are simultaneously used to identify the remaining parameters of the model. However, it is instructive to examine some comparative statics to see how the different parameters determine different observable aggregates. The parameters θ and τ both play a role in determining the degree of sorting in the model. A larger θ will tend to reduce the extent to which more educated households are matched with more productive regions, as idiosyncratic preferences will be more important in deciding where the household lives.

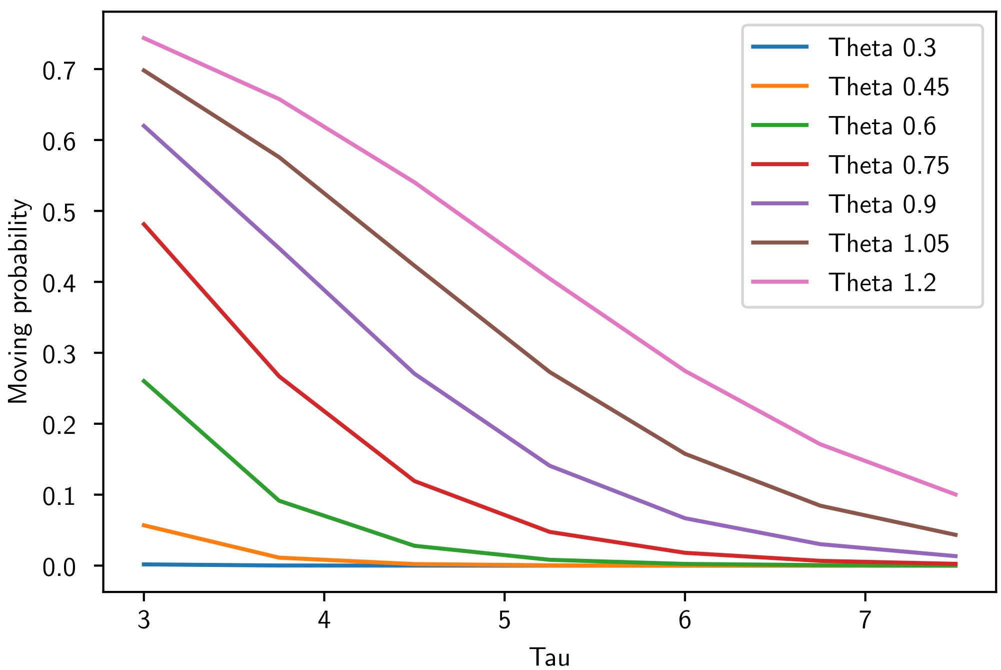
In my model each period each household chooses a location j' conditional on their current state, that is their age, bond wealth, education, and current location. I observe the household's location choice, j', as well as their age, education and current location. The remaining state, the households (non-housing) financial net worth is therefore unobserved.

By contrast a larger moving cost, τ increases the degree of sorting of households along education and city productivity lines. A larger moving cost increases the costs of moving to a city to take advantage of an idiosyncratic preference shock. With a larger moving cost, a household will need a larger idiosyncratic preference shock to induce them to move. Thus, in a sense larger moving costs reduce the relative importance of idiosyncratic preference shocks and hence increase the degree of sorting.

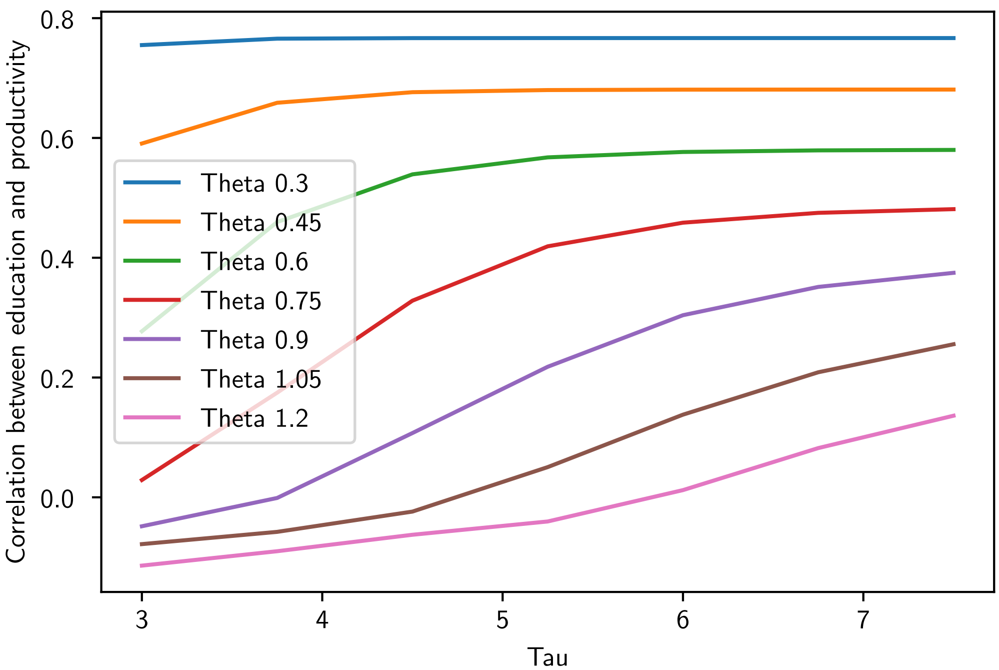
Both θ and τ also impact the moving rates. A higher τ increases the utility loss the household receives when they move between two regions and so decreases the moving rate. A higher θ increases the moving rate. This is because a larger θ increases the size of the i.i.d. preference shocks. Since these shocks become relatively more important the household is more likely to move to a new region to take advantage of their idiosyncratic preference for living in a region.

In Figure 1.9 I plot the mean moving rate against values of τ for different θ. As one can see a larger θ increases the moving rate for a given value τ and that the mean moving rate is decreasing in τ. To measure the degree of sorting along productivity education lines I measure the correlation between the city-level productivity and the average education level. I plot this correlation in Figure 1.10.

1.9. Moving probability comparative statics.



1.10. Education city productivity correlation comparative statics.



1.6.5.2 Identification and Estimation

Owing to the structural assumption on the household's preference shocks the CCP of selecting a particular region takes the tractable form of equation 1.10. Since households die deterministically these CCPs can be computed by backwards recursion.

In my model each period each households choose a location j' conditional on their current state, that is their age, bond wealth, education, and current location. I observe the household's location choice, j', as well as their age, education and current location. The remaining state, the households (non-housing) financial net worth is unobserved in my data. Denote the individual state excluding their risk-free assets as x_i and the bond wealth as b_i. Households thus choose a location j_i' conditional on observed x_i and b_i.

Denote the household's utility parameters that are to be estimated as Ξ. Let $\pi(j_i'|x_i, b_i; \Xi)$ be the probability that a households chooses location j_i' conditional on state x_i and b_i, given model parameters Ξ. Label the probability distribution of household bond wealth conditional on state x_i as $\phi(b|x_i; \Xi)$ [23]. Then by the law of total probability I can write the likelihood that a household with observed state x_i who chooses location j_i' as

$$\pi(j_i'|x_i; \Xi) = \sum_{b \in Supp(B)} \pi(j_i'|x_i, b; \Xi)\phi(b|x_i; \Xi), \tag{1.29}$$

where the CCP, $\pi(j_i'|x_i, b; \Xi)$, comes directly solving the households problem, $Supp(B)$

[23] The model implied distribution of the risk-free will be discrete due to the assumption that all households start with no financial wealth.

is the support of bond wealth given the state x_i and parameters Ξ. Given observations N of household location choice, conditional on age, education, and previous location, the maximum likelihood estimate of the parameters solves,

$$(\hat{\Xi}, \hat{\phi}) = \arg\max_{\Xi,\phi} \sum_{i=1}^{N} \log(\pi(j_i'|x_i;\Xi)) = \sum_{i=1}^{N} \sum_{b} \log(\pi(j_i'|x_i, b; \Xi)\phi(b|x_i;\Xi)). \quad (1.30)$$

As (1.30) is not additively separable estimating Ξ requires an knowledge of ϕ and the distribution of ϕ depends on the estimates Ξ. To estimate this I therefore use the expectation maximization method proposed in Arcidiacono and Miller (2011). Using the ECCP algorithm requires the following assumption:

Assumption 1. *Every household's individual state follow a Markov process that depends only on the previous realizations of the the household's individual state.*

It is easy to see that states in my model are Markovian. In the model the household's age, education are deterministic and depend on only the previous values of age and education respectively and hence are Markovian. The model implies policy functions for bond wealth, $b'(a, b, e, j, j')$, and location choice that depend only on current state and choice. Thus, in the model bond wealth and current location are also Markovian in current state.

Finally, in estimation I assume that 1990 represents a steady-state for the US economy and estimate the model to the 1990 data [24]. The 5% Census microdata sample Ruggles et al. (2021) contains the location choices of individual households

[24] This is of course a strong assumption. However, given the difficulty in computing an out-of-steady equilibrium, and the need for repeated computation in estimation I defer further discussion to future research. See Ahlfeldt et al. (2020) for a more detailed discussion of transitional dynamics in models with migration costs.

conditional on their age, education and previous location. Table 1.1 contains summary statistics for this data.

1.6.5.3 Interpretation of Estimates

The results of the estimation are in table 1.8. In line with theory all parameter estimates are positive and significant. In isolation the estimated parameters are difficult to interpret. To ease the interpretation I calculate a first order approximation to the compensating variation when the parameter is change for the hypothetical mean household. I define this mean household as a household with the mean consumption of households across age, education, and location[25]. I then interpret the parameters in dollar terms by calculating the change in utility, $\Delta U(x)$, as a result of a change in parameter x, and calculating change in consumption required to make the household indifferent to this change. This is given by

$$\Delta c u'(c) \approx \Delta U(x). \tag{1.31}$$

The Δc term then is an approximate change in consumption required to keep the household indifferent.

From table 1.8 we see that the estimated parameter ω is 0.0027. This implies that the average worker would be indifferent between a one standard deviation increase in amenities and a 0.43% increase in consumption or approximately \$165 in 1990 terms. However, the concave nature of consumption utility mean that this differs across households. The most educated households would be indifferent between

[25] I exclude the youngest cohort as they have no moving costs by assumption

$263 increase in consumption and a one standard deviation increase in local amenity supply.

1.8. Estimated households utility parameters.

Parameter	Estimate	Std. Err.	z	P> |z|	[0.025	0.975]
ω	0.0027	0.00019	14.102	0.000	0.0031	0.0023
θ	0.4937	0.00169	291.981	0.000	0.4969	0.4904
τ	3.0998	0.01098	282.322	0.000	3.121	3.078

In line with the literature, Kennan and Walker (2011), the estimated intercity moving costs are large. The estimated value of τ is 3.1. This is equivalent to $118,577 for the average household in the model. A large migration cost is to be expected in this model given the only 11.7% percent of households report moving cities in the past five years.

The final parameter that needs to be interpreted is θ, which governs the importance of preference shocks to the household. To interpret this parameter, I use equation 1.11, the expected utility prior to the realization of the preference shock. On average for a household at location j who remains in location j they would be indifferent between a $12,922 increase in yearly consumption and not receiving preference shocks.

For households that receive a preference shock large enough to induce them to move they would expect to be indifferent between an increase in consumption of $96,869 and not receiving the shock. While this seems large, given how large the estimated moving costs are it is no surprise that households need a large preference shock to induce them to move [26].

1.6.6 Evaluating the Model

Here I explore some steady-state results from the model. I solve for the steady-state of the model in 1990 using the estimated parameters. I then use this to derive some comparisons. In the model we have that the mean moving rate across all households is 10.12%, which is somewhat smaller than the 11.7% observed in the data. [27]

I calculate the correlation between city-level productivity and the the 1990 share of the population with at least a bachelors degree in the data and in the model. Roughly speaking the stronger this correlation is the greater the degree of spatial sorting along education and city productivity lines. The city-level productivity in both the model and the data is calculated the city-level fixed effect from the regression in (1.23). In the data I find a positive correlation of 0.32 between the share of the population with a college degree and the city's labor productivity. In the 1990 steady-state of the model the same correlation is 0.62, or about twice as large as the data. Thus, the model predicts more sorting along education and city productivity lines than we see in the data.

As discussed in section 1.3 in the data there is strong correlation between the return on housing in the city and the productivity shock. In Table 1.9 I show the

---

[26] One could of course view the sum of the preference shock and the moving costs as a single random variable where the mean differs depending whether the household currently lives in the city.

[27] Recall that we in effect choose the parameters of the model to match *all* the moments of the data and so we don't match the moving rate exactly.

results of a regression of the log change of the region's house price between 1990 and 2019 on the shift-share productivity shock for the city during the same period. I perform this regression using both the observed changes in regional house prices and 1990 and 2019 steady-state prices from the model. In both cases the coefficients are positive and highly significant. While this regression is purely to illustrate the observed correlation, in the data a 1% increase in the city productivity shock increases the local real house price by 4.7% over the 1990 to 2019 period. In the model a 1% increase in productivity over this period increases the steady-state price by 7.4%.

1.9. Regression of log change in real house prices on log productivity shock.

Dependent Variable:	Δ Log House Price	
	Data	Model
Model:	(1)	(2)
Variables		
(Intercept)	-2.734**	-4.841***
	(1.242)	(1.042)
Δ Log Productivity	4.655**	7.366***
	(1.808)	(1.515)
Fit statistics		
Observations	71	71
R^2	0.06112	0.31539
Adjusted R^2	0.04751	0.30547

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

1.7 Counterfactuals

Using the estimated model, I can evaluate the heterogeneous effects of several counterfactuals on different cohorts. Increases in housing elasticities make high-productivity and high amenity cities more affordable. In steady-state consumption and amenity utility increases more for young than for older cohorts. In the short-run however, older households may see the values of their homes decrease relative to the world where the housing supply is restricted.

In addition to changing the land-use regulatory environment I explore whether other changes can generate similar changes in consumption. During the COVID-19 the fraction of workers who work from home increased dramatically, motivated by this I examine the long-term impact that this change would have consumption across different household groups. Finally, I examine how creating new urban areas impacts different household groups.

1.7.1 long-run Impact of Removing Land-use Regulations

To assess the long-run impact of changes in land-use regulations and other policies I compare the steady-state of the counterfactual economy to the steady-state of the baseline economy. In particular I compare how consumption and the utility from amenities changes for different households between these different steady-states.

I compare the 2019 steady-state with the estimated local housing supply elasticities to one where I remove all land-use regulations. Since the land-use regulation measure has real support, this implies that $\alpha_j = 0$ for all cities j[28]. This implies that

[28]In practice the smallest value of α_j that I observe in the baseline economy is 0.22. Deregulating

the local housing supply is infinitely elastic. When supply elasticity is infinity it is easy to see that the price of housing is equalized in all regions. In both steady-states city amenities, city productivity levels, and total population measure are set to their 2019 level.

There are three potential sources of changes in household well-being when comparing these two steady-states. First consumption may increase as households move to higher productivity regions. Intuitively local area productivity will be positively correlated with house prices, as households will prefer areas with high-productivity to those with low productivity, which raises local housing demand. Thus, when these elasticities rise, households will tend to move to regions with higher productivity, increasing their income. Secondly, housing costs will decrease within a region. This increases the after housing costs income of households within the city and hence their consumption. Finally, since regions with good amenities have expensive housing, increasing elasticities will tend to increase household amenity utility, are able to afford high these high amenity supply cities.

In aggregate I find that removing all land-use restrictions increases aggregate consumption by 7.13%. Steady-state incomes rises by 3.34%. The remaining increase in aggregate consumption is due to decreased housing costs. Income dispersion between cities, measured by the coefficient in variation in incomes between them, declines by 22.09%, as highly productive cities now have a greater number of less educated workers. In total amenity utility grows by 8.55% between these two steady-states as workers move from low amenity locations to high amenity areas.

In Figure 1.11 I plot how the equilibrium consumption changes for different house-

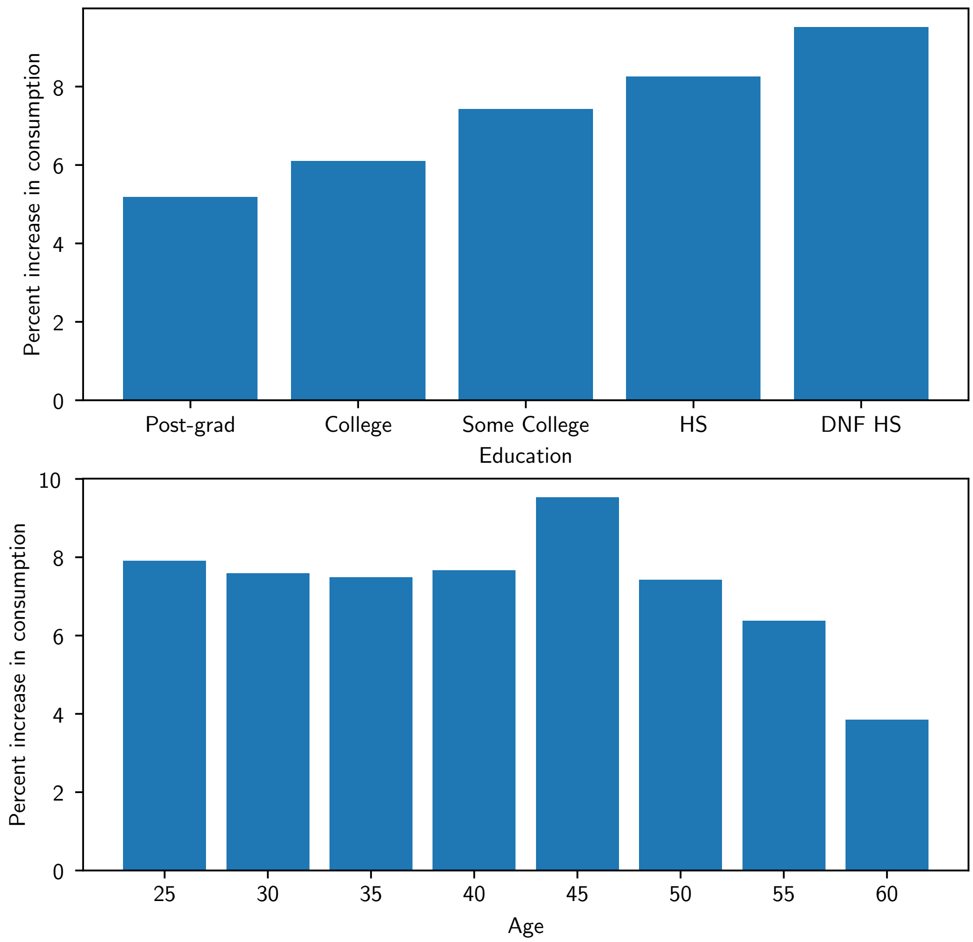
all cities to this level leads to a rise in aggregate consumption of 6.59% versus 7.13% when all regulations are abolished. The other results are also quantitatively similar.

hold groups from the baseline 2019 steady-state to a steady-state where land-use regulations are abolished. The top portion of Figure 1.11 plots the percent change in consumption between the two steady-states conditional on household education. From the figure we can see that consumption growth between the steady-states is largest for households with the lowest level of education and that consumption growth monotonically decreases as household education increases. The rise in consumption for the least educated group is over 9.5%, while the percent increase in consumption for the most educated group is 5.1%. Thus, removing land-use regulations leads to a reduction in consumption inequality between household groups.

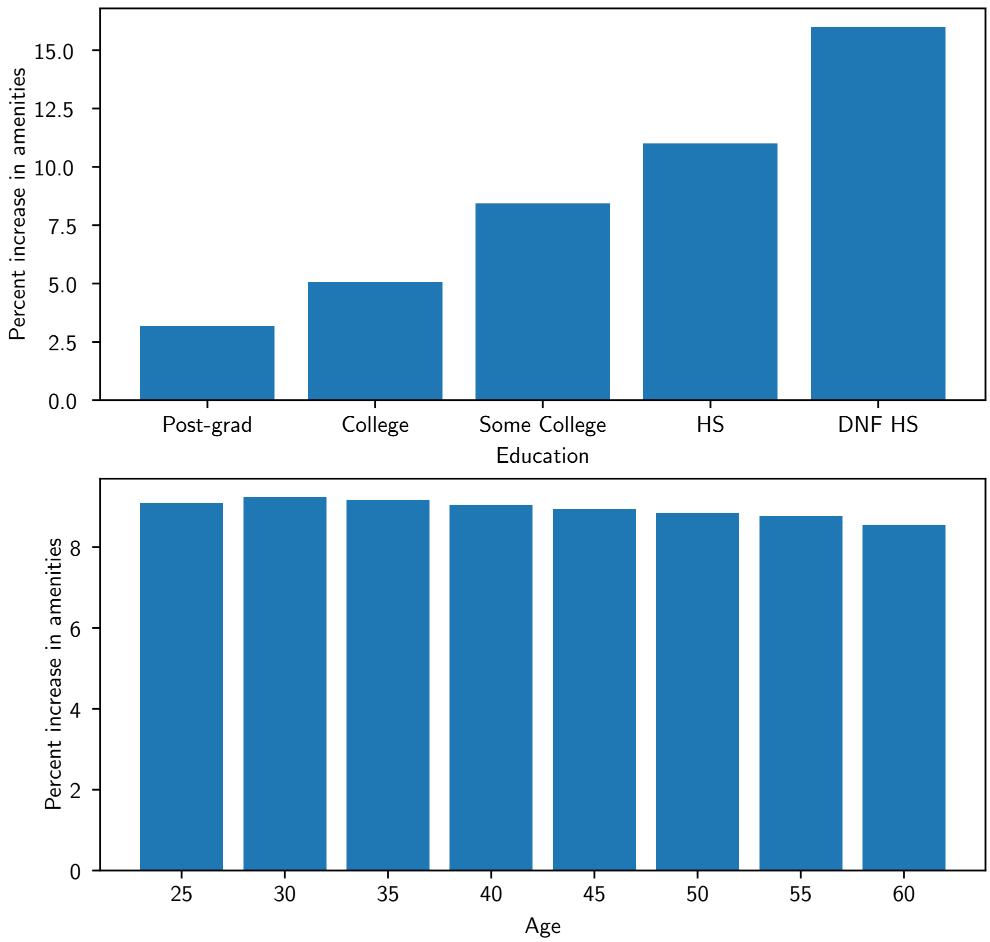
The bottom portion of Figure 1.11 plots the percentage change in consumption for households conditional on age between these two steady-states. The oldest households see the smallest gain in consumption when we compare these two steady-states, with a consumption gain of approximately 3.8%, or less than half that of the youngest households. Interestingly households aged 45 are the ones with the largest gains in consumption. Their consumption grows by 9.5% between the two steady-states. This is an artifact of the life-cycle pattern of housing ownership and the household borrowing constraint. When a household enters the period in their life where they become a homeowner the borrowing constraint incentivizes them to move from being a renter in a high-cost region to a homeowner in a lower cost region or to significantly reduce their consumption that period. When housing prices are equalized across regions the incentives to move disappear and the reduction in prices means the borrowing constraint becomes less binding.

In Figure 1.12 I plot the growth in amenity consumption between the two steady-states conditional on education and age. In the upper portion of Figure 1.12 I plot the growth in amenity utility between the two steady-states conditional on

1.11. Steady-state consumption changes when land-use regulations are removed.



1.12. Steady-state amenity changes when land-use regulations are removed.



education. This figure shows that there is a strong relationship between household education and the increase in amenities between the two steady-states. The least educated households see the largest gains in amenities, with an increase in amenity consumption of about 16%. In comparison the most educated household group see a rise in amenity utility of just 3.19% between the two steady-states. Thus, removing land-use restrictions reduces amenity inequality between households.

The lower section of 1.12 plots how amenity consumption grows between the steady-states conditional on age. Unlike consumption there is only a weak relationship between age and amenity consumption growth between the two steady-states. All age cohorts receive similar gains in amenity consumption, due to the persistence in location choice over the household's life and the imperfect correlation between amenities and productivity. The oldest age group see the smallest increase in amenity consumption with a rise of approximately 8.5%. The growth in amenity utility is 9.24% for the 30 year old group.

1.7.2 short-run Impact of Removing Land-use Regulations

In steady-state I showed that relaxing land-use regulations increased consumption of all education and age cohorts. In reality there is substantial opposition to new development in many parts of the country. Not in My Backyard (NIMBY) organisations exists on all sides of the political spectrum and have succeeded in blocking new construction in their area, citing disruption caused by construction and congestion externalities due to increased population. However, it is difficult to separate concerns about negative externalities, from the substantial economic benefits that existing homeowners receive when they limit the supply of new housing.

Existing homeowners reap substantial capital gains on their homes when they limit the supply of new housing. Older homeowners purchased their homes in an era where land-use regulations were less pervasive and hence the housing supply was much more elastic. When local homeowners vote for stricter land-use regulations in their city, they increase the marginal cost of producing housing and hence increase its price. As housing is the largest asset in the portfolios of most households the life-time consumption benefits from these capital gains can be substantial. These gains come at the expense of younger households, as they have no housing wealth and must purchase a home to live in [29].

To quantify the winners and losers from land-use regulation I examine a transition between steady-states with the baseline values of land-use regulations and when land-use regulations are abolished. In these transitions I begin at the 1990 steady-state. I assume that changes to local amenity supply, national population, and productivity are completely unanticipated, but fully known upon impact. Local amenity supply is the supply in the region at each time. Productivity is the 1990 productivity plus the shift-share productivity shock up until that period. Specifically, to compute the productivity growth I use the cities 1990 wage bill shares across all 3-digit level industries and multiply that by the national change in wages for that industry excluding the city itself. Summing across all industries yields the city's change in productivity. The construction of local productivity is detailed in section 1.5. I assume that 70 years after the last shock we reach the new steady-state[30]. The

[29] It is possible that rent controls could impose further costs on the youngest cohorts. Diamond et al. (2019) finds that rent control reduced the supply of rental housing in San Francisco and hence increased rents for new households, while protecting incumbents.

[30] The curse of dimensionality significantly magnifies the difficulties that arise with computing the wealth distribution along a transition compared with the steady-state case, as the dependence

total population will equal the total US population in the cities in my study.

To compute the equilibrium transition prices I compute a series of temporary equilibria and iterate until the expectations used in the temporary equilibria converge to realized equilibrium prices. To do this I first create a guess a sequence of prices for each city. Then working forward in time, I compute a temporary equilibrium, that is I compute the equilibrium price vector at the date, given the past computed equilibrium prices and the future price guess. Once I have computed all temporary equilibria, I restart computing the temporary equilibria beginning again in the first period and working forward, now using the previous sequence of temporary equilibria as the price guess. I continue to iterate until the differences in temporary equilibrium prices is small[31]. A full explanation of the solution method is in Appendix 1.B.

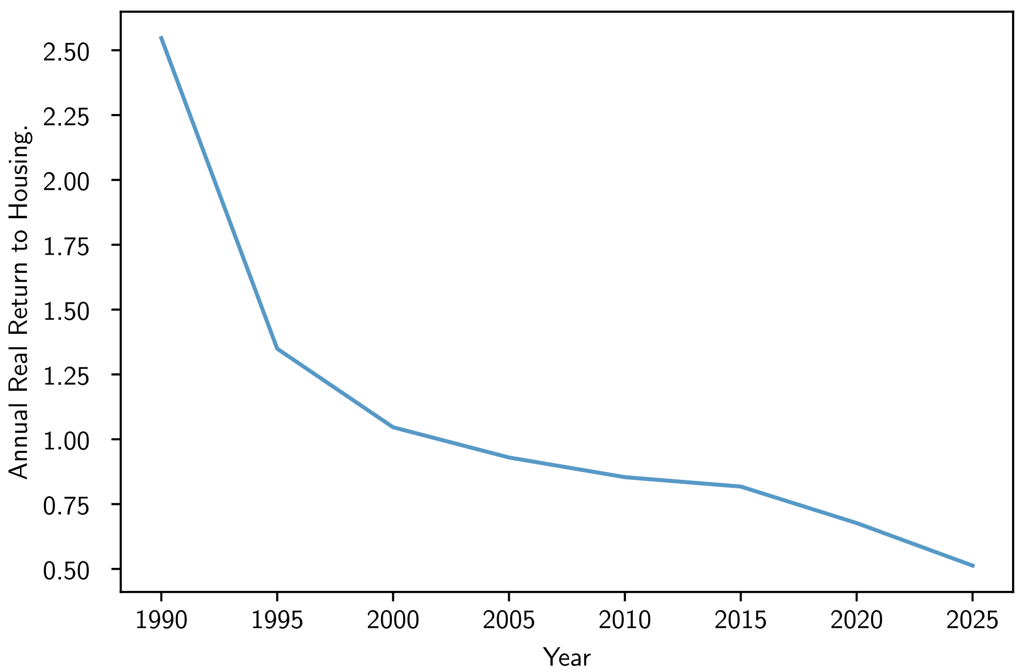
I first examine how the return on housing evolves along the transition. In Figure 1.13 I plot the 1990 population weighted annualized return to owner-occupied housing during the transition. This return is simply the capital gain on a home in the region. As we can see homeowners experienced substantial unanticipated appreciation on their homes, increasing their net worth. After 2020 there are no further shocks to the economy, and I allow the economy to gradually approach the new steady-state. This Figure shows that homeowners, particularly the oldest generations, experienced substantial anticipated capital gains on their housing.

To calculate which cohorts are winners and which are losers I calculate the present discounted mean consumption of different generations along both the actual and the

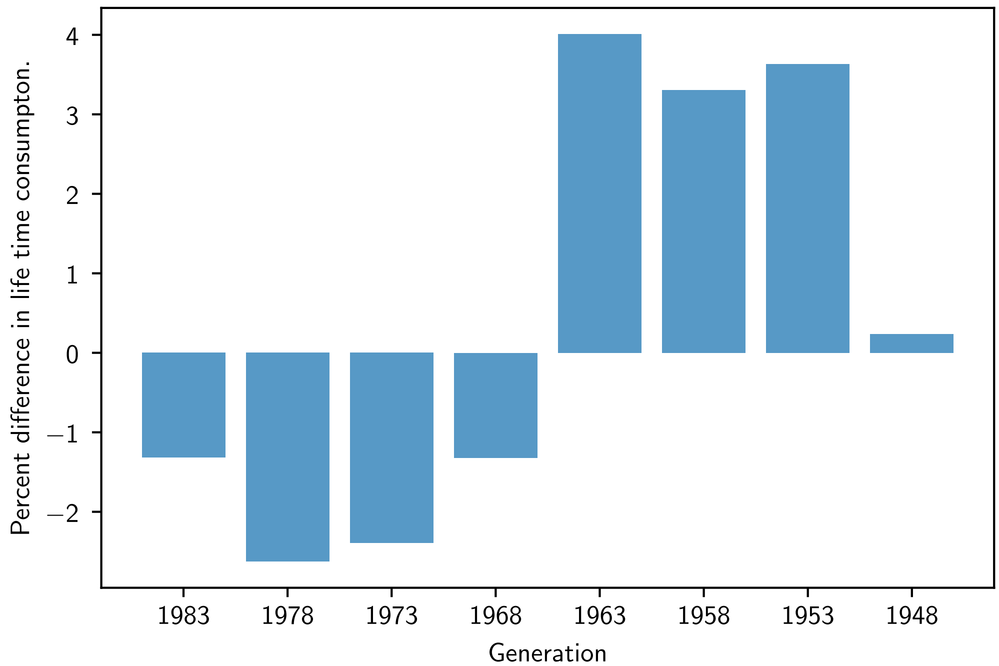
on time increases the state-space considerably.

[31]In these transitions assume that households fully anticipated the entire path of the changes to productivity, amenities etc. upon impact. The actual expectations in 1990 may not match this exactly. However, alternative methods for forming expectations can easily be incorporated into my framework. I leave this as something for further research.

1.13. Model annual real return to owner-occupied housing.



1.14. Consumption changes along transaction.



alternative transition. I then plot the percentage difference in household consumption between these two transitions in Figure 1.14. As we can see from Figure for cohorts born before mid-1960s their life-time consumption is larger when land-use regulations are equal to their baseline level. This is because the capital gains they receive on their housing more than offsets the fact that relative to the counterfactual households live in less productive cities.

From Figure 1.14 we can see that for younger cohorts and particularly those born after the 1970s their consumption is larger in the counterfactual transition where I remove all land-use regulations. This is because these groups are largely renters when the shock occurs and hence don't receive most of the capital gains on housing. In the counterfactual world these households will tend to locate in higher productivity cities, which increases their wages. Furthermore, when they purchase their home they benefit directly from cheaper housing, increasing their consumption.

1.7.3 Switch to Working from Home

While the direct influence of policy makers on the ability of workers to work from home is limited, the COVID-19 pandemic has exposed that workers can increasingly work remotely, Brynjolfsson et al. (2020). While it is not possible to accurately forecast any long-term changes in work arrangements, it seems likely that an increased fraction of households will work from home rather than commute, Barrero et al. (2021), Gupta et al. (2021) [32]. It is thus interesting to consider how the reallocation across space of workers due to a rising fraction of households working from home will change welfare.

When a worker's productivity is no longer tied to their location, they may relocate to take advantage of better amenities or reduced housing costs. Anecdotally many workers, who were now working from home, left the San Francisco Bay area during the COVID-19 pandemic to take advantage of lower housing costs in other regions Bowles (2021). Workers leaving a city when they can work from home causes housing costs to fall, increasing the after-housing cost income of workers who remain [33]. In

[32] As the model is estimated so that a single period represents five years it is ill-suited to examining the very short-term impact of households leaving cities during the COVID-19 pandemic.

[33] There are several difficulties that arise in quantifying the short-run impact of the COVID-19 pandemic on different housing markets. First, since many of these high-cost cities, such as New York, were most affected by the initial pandemic waves, the expected duration of the pandemic would play a crucial role in determining the movement of workers and prices. Moreover, many

steady-state, new workers, who by assumption cannot work from home, may enter the city now that prices are lower, and hence increase the productivity that's available to them.

A rising fraction of households working from home is not Pareto improving, however. Workers who don't work from home and are living in low-cost cities that experience an inflow of workers will be worse off if they stay. This is because in steady-state housing costs will increase and their wages will not change. This is similar to what happened in many smaller low cost cities, such as Boise, during the COVID-19 pandemic [34].

To model a work from home policy I follow the literature and assume that 20% of workers can now work remotely. I assume that a worker who works from home receives the national average productivity for their education type from the previous 2019 steady-state. That will be the weighted average city-level productivity, with weights that depend on the measure of type e in each city. I then solve for the new steady-state of the model, which will show the long-run effect of a shift to working from home. This means that the productivity of a household who is working remotely may increase or decrease. For those living in a low productivity region it will rise, while those in high-productivity regions it will decrease.

I find that aggregate consumption grows by a modest 1.39% across all workers. Workers who don't work remotely see a rise of 1.12%, while workers who work from home see a rise in their consumption of 2.42%, implying that the workers who benefit the most from the change in the work environment are those that can work from

---

high-cost high-amenity regions had more stringent lock-down policies, reducing amenity supply to a greater extent than in other cities. This would further reduce housing demand for these regions.

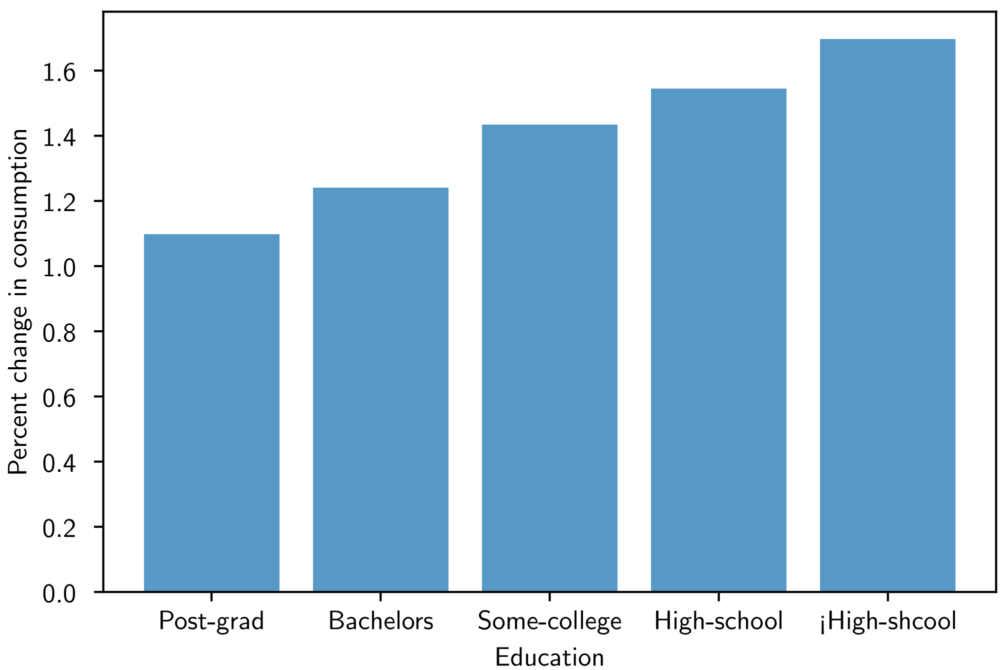
[34] See https://www.apartmentlist.com/research/national-rent-data for trends in city rent data.

home. However, the model does not encompass all the benefits of the shift to working from home. Commuting times will fall, as many workers no longer need to commute and workers who don't work remotely face less congestion.

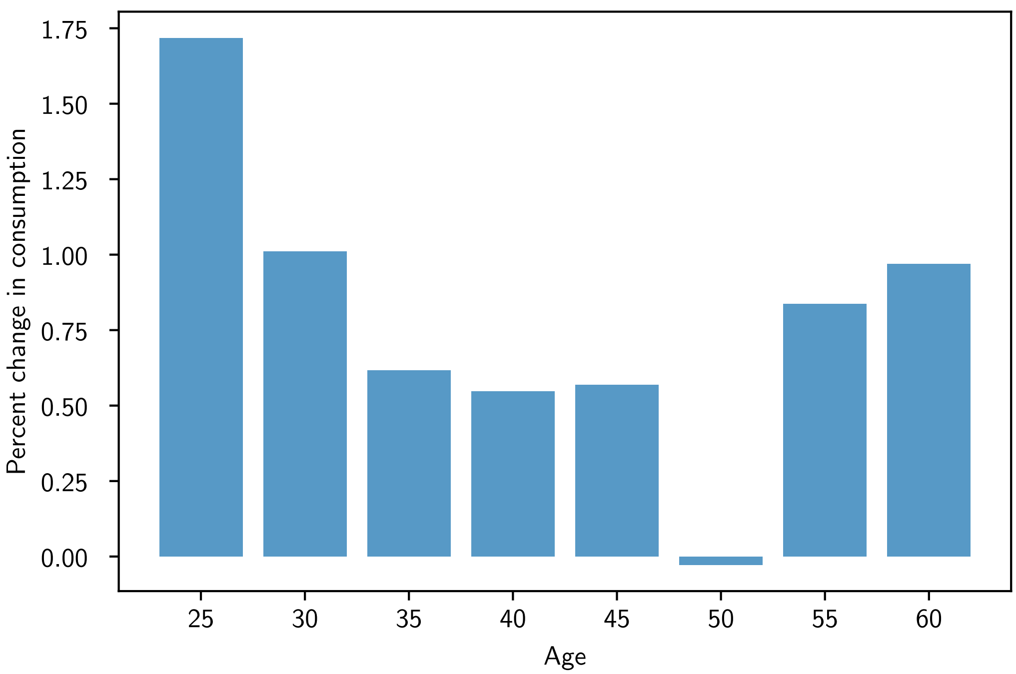
In Figure 1.15 I show that the least educated workers see the largest increases in consumption of workers. This Figure plots how consumption changes between the baseline 2019 steady-state and a steady-state where 20% of workers work from home. From the Figure we can see that on average households with a post-graduate education see consumption gains of about 1.1%, while those with less than high school education see gains of about 1.7%. Recall that I assumed that all workers are affected by the switch to working from home equally. So, this result is not due to difference in the propensity to work remotely but rather due to differences in the likelihood of reallocating across cities and variation in spending on housing across households.

I find that the change in consumption caused by the switch to remote working is non-monotonic in age and approximately U-shaped. In Figure 1.16 I show that while the youngest cohorts see the largest increase in consumption, middle-aged households see almost no rise in their consumption and the oldest cohorts see a moderate rise in their consumption. The youngest group see an increase in consumption of 1.72% while the 50 year old group see approximately no change in consumption. This non-monotonic relationship arises as the workers productivity is hump shaped in their age. Thus, as young workers approach middle age, they tend to relocate to higher productivity places. As a worker continues to age their individual level productivity begins to fall and so they move to lower cost, less productive places. When working from home this incentive to move between regions of different productivity disappears and so we observe this non-monotonic relationship.

1.15. Consumption change when 20% of workers work-from-home by education.



1.16. Consumption change when 20% of workers work-from-home by age.



1.7.4 Creation of New Cities

Despite recent progress in reforming land-use regulations in states such as California [35], it seems unlikely that there will be a complete removal of all land-use regulations in the near future. An alternative to deregulating zoning in existing cities could be the creation of new cities on unincorporated land. In states such as California incorporating new cities was used to increase the housing supply in the early post-war era. In the 1960s 46 cities were incorporated in the state, while there have been no new incorporated cities in the past decade.

To simulate the impact of creating a new urban area I increase the number of urban areas in the model by 10%. I assume that each of these new cities has the population weighted mean productivity, amenities, and housing supply elasticity from the 2019 steady-state[36]. I then solve for the steady-state of the model. The creation of new cities in the model mirrors how new urban areas were created and existing ones saw the development of new suburban communities in the post-war era.

I find that a 10% increase in the number of urban areas leads to a 1.41% increase in aggregate consumption compared to the 2019 baseline steady-state. The rise in aggregate consumption is exclusively due to the reduction in housing costs. In fact average incomes are approximately unchanged between the two steady-states. In the model all else equal adding a new region will decrease the probability that a household selects from the existing places. Thus, in equilibrium the creation of

[35] With the signing into law of state level legislation S.B. 9 and S.B. 11 in September 2021 state law will now override many local zoning restrictions

[36] The magnitude of the changes that creating new cities induces is sensitive to the assumed amenity supply and productivity of the new cities. The more productive the new regions the larger the magnitudes.

new cities reduces demand for all existing regions, decreasing housing costs in these regions. This in turn increases the after-housing costs income of households and hence their consumption.

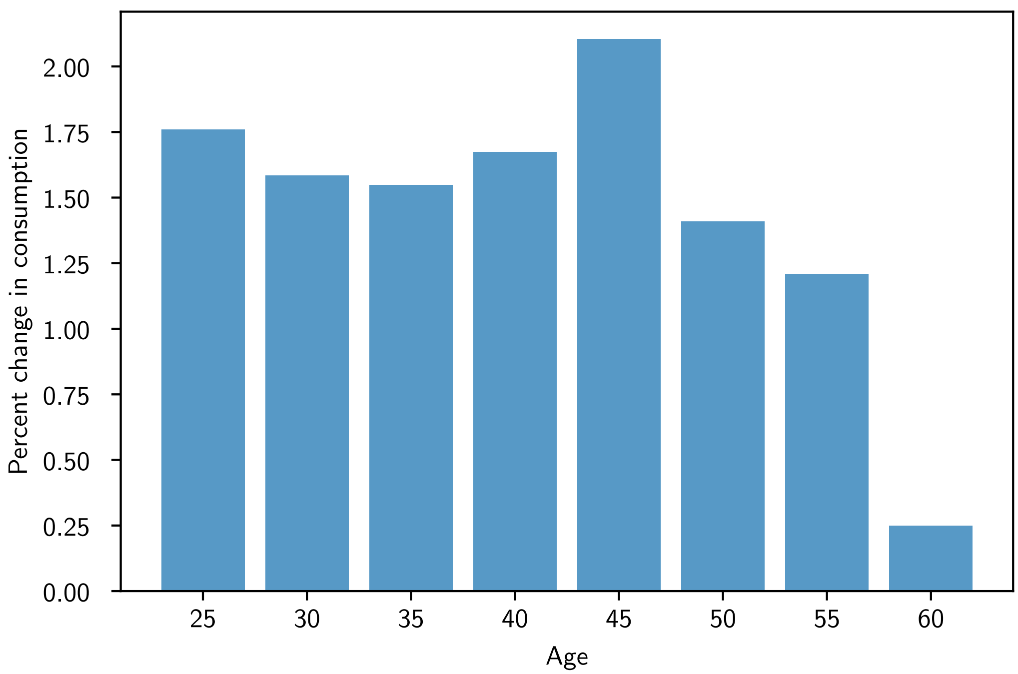
The effect of creating new cities is also heterogeneous across households. In Figure 1.17 I plot the how consumption changes conditional on the age of the household. From the Figure it is clear that younger households see larger gains in consumption than older households. Consumption for the youngest households rises by 1.76% while for the oldest household groups it grows by only 0.02%. Again, the largest gain in consumption is when the household becomes a homeowner, when consumption grows by a little over 2.1%.

Adding new cities also reduces consumption inequality conditional on education. In Figure 1.18 I plot the consumption growth between the baseline 2019 steady-state and one where the number of cities is increased by 10%. As one can see from the Figure the percentage increase in consumption is more than twice as large for the least educated households compared with the most educated ones. Households with less than high-school education see increases in consumption of 2.05%, while those with a post-graduate degree see increases in consumption of 0.89%. This differences in the consumption gains are due to differences in expenditure shares on housing for households with different education levels and because low-wage workers are able to afford more productive regions, as prices fall.

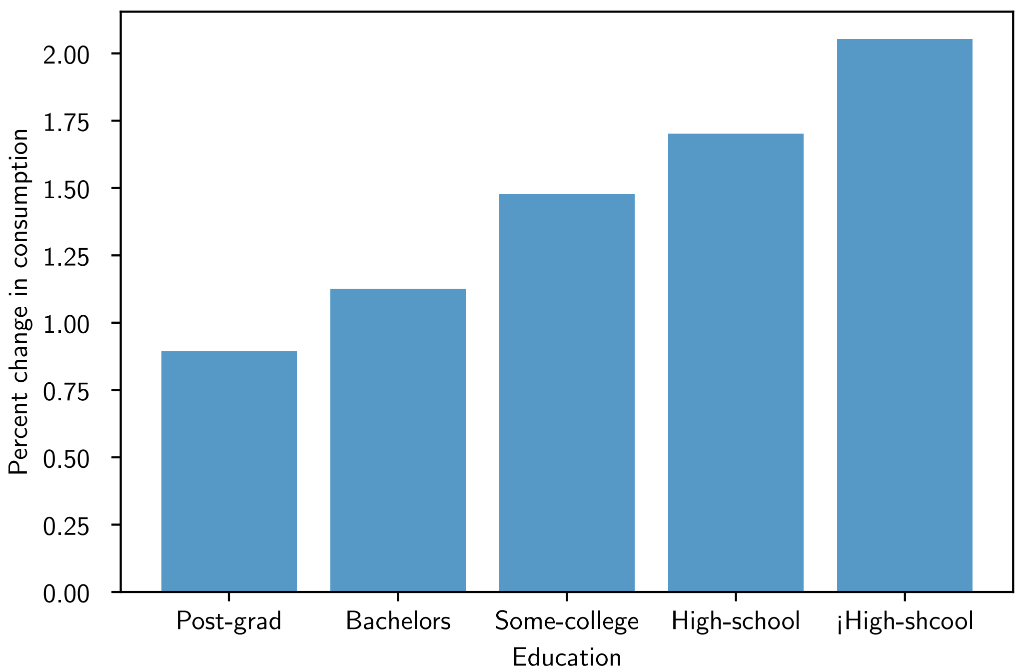
1.8 Conclusion

Pervasive land-use regulations, such as setback requirements, urban growth boundaries, and height limits, have had an enormous effect on the US economy. By increas-

1.17. Consumption change with additional cities by age.



1.18. Consumption change with additional cities by education.



ing housing costs in the most productive cities, they have allocated labor away from these regions towards less productive but cheaper areas. The costs of land-use regulations potentially vary across households due to differences in portfolio composition and their effect on the spatial sorting of households.

My paper quantifies the heterogeneous costs of land-use regulations across households of different ages and education levels. To quantify these costs, I estimate a structural dynamic spatial equilibrium model of household location choice. My model features overlapping generations of households who differ in terms of age, education, wealth, and current location. Households choose a sequence of locations to own or rent housing, trading off local amenities, wages, moving costs, and housing costs. To quantify the effects of land-use regulations, I conduct two sets of counterfactual experiments exploring the long-term and short-term impact of land-use regulations.

To examine the long-run effects of land-use regulations across different household cohorts, I compare steady-states with the estimated land-use regulations to one where all land-use regulations are abolished. In my model removing land-use regulations makes the local housing supply infinitely elastic, equalizing prices across cities. Abolishing land-use regulations increases aggregate consumption by a modest but significant 7.1%, as workers move from less productive areas to the most productive cities. However, abolishing land-use regulations reduces the degree of spatial sorting and hence this effect differs across households depending on their education and age. The least educated households see gains in consumption that are about twice as large as the most educated, and the youngest households see consumption gains that are about three times as large as the oldest cohorts.

My finding that all cohorts benefit from removing land-use regulations raises the question as to why they are so pervasive. To answer this, I examine the transitional dynamics between steady-states. I begin at the 1990 steady-state and feed in the estimated sequence of productivity shocks, amenity supply and total population. I compare how consumption of different generations changes along a transition with the estimated elasticities and one where land-use regulations are abolished. I find that removing land-use regulation hurts cohorts born before the 1960s, while benefiting younger cohorts. This is because the large capital gains that older homeowners receive on their home more than offsets any losses from living in less productive regions.

The results of this paper suggest that while changes in land-use regulations can bring about substantial benefits, particularly to lower-income households, in the short-run many older cohorts will be negatively affected. Since older households tend to be more politically active, it is no surprise that even relatively modest reform proposals, such as S.B. 9 and S.B. 10 in California, face substantial opposition. Motivated by these difficulties, I explore whether allowing workers to work from home or increasing the number of cities can substitute for reducing land-use regulations. My results imply that while the benefits of these alternative policies are qualitatively similar, quantitatively the benefits are only 20% as large as removing land-use regulations from existing urban areas.

The current study can be extended in several directions; an important one is noted. For tractability, this study does not feature any local agglomeration economies. If agglomeration economies proved significant, then strict land-use regulations may decrease long-term total productivity growth. Potentially contributing to the slowdown in total factor productivity growth over the past 30 years.

APPENDIX

1.A Computing the Steady-State

I first create a non-linear grid for bonds around zero, so that the grid density is higher close to zero. All other state variables are discrete and can be exactly represented on a finite grid.

To compute the steady-state of the economy we I begin with an initial guess for the price of housing in all markets. Conditional on this price guess I can solve the household's value function, working backwards from the final period of life until birth. The next step is to solve forward for the measure of households in each city conditional on the price guess. Given the households value function conditional on choosing j' given its state variables, $v(a, b, e, j, j')$, and equation 1.10 we have the probability of choosing j' conditional on state variables.

Recall that by assumption in the initial period there are no moving costs so the value function does not depend on j. I also assume that the households starts life with no bond-wealth. Hence choosing an arbitrary initial location, l, the measure of households in each city j' in period zero of their life [37] is

$$\mu(a, e, j') = \frac{\exp(v(a, 0, e, l, j'; S))^{\frac{1}{\theta}}}{\sum_k \exp(v(a, 0, e, l, k))^{\frac{1}{\theta}}} \tag{1.32}$$

which we compute by interpolating over b and the measure in city j' conditional on type e and age a is written as $\mu(a, e, j')$. We of course set $a = 0$ for the first period of life.

[37] I index beginning with zero not one.

We now need to solve for bond wealth for all future periods of life. This will be an exact function of all previous locations, age, and education. As previously argued, since the number of locations is large it is not possible to compute this exact function. I therefore calculate bond wealth conditional on current location. Let $b(a, e, l, j')$ be bond wealth conditional on age, education, previous location, and current location. For age one we then have that

$$b(1, e, j, j') = b'(0, 0, e, l, j) \tag{1.33}$$

where we use the policy function for b' at age zero, education e, initial location l, interpolating over zero bond wealth. The next step is to calculate bond wealth is to use the probabilities and measures to get bonds conditional on current location and not past cities. Define a conditional probability $\tilde{\pi}(a, e, k, j')$ as follows,

$$\tilde{\pi}(a, e, k, j') = \frac{\pi(a, b(a, e, k, j'), e, k, j') \times \mu(a, e, k)}{\sum_l \pi(a, b(a, e, l, j'), e, l, j') \times \mu(a, e, l)} \tag{1.34}$$

where we are interpolating over b in our function $\pi(a, b(a, e, k, j'), e, k, j')$. I then use this to calculate the bond wealth conditional on current location as follows

$$b(a, e, j') = \sum_k \tilde{\pi}(a, e, k, j') \times b(a, e, k, j') \tag{1.35}$$

I calculate bond wealth conditional on previous location j and current location j' for all ages after one using

$$\zeta(a,e,l,j) = \frac{\pi(a,b(a,e,l),e,l,j) \times \mu(a-1,e,l)}{\sum_k \pi(a,b(a,e,k),e,k,j) \times \mu(a-1,e,k)} \tag{1.36}$$

$$b(a+1,e,j,j') = \sum_l b'(a,b(a,e,l),e,l,j) \times \zeta(a,e,l,j) \tag{1.37}$$

We can then iterate forward using the above steps and calculate $\mu(a,e,j')$ for ages and types. Then using the data measure of households with education e, $g(e)$, and the fact that life last $\overline{a}$ periods the measure in each city equals

$$\mu(j') = \sum a,e(1/\overline{a}) \times f(e) \times \mu(a,e,j') \tag{1.38}$$

where I use the implicitly assumed independence between a and e. [38] With the measure in each city I then use 1.19 and the measure to create a price implied by housing supply conditional on the level of demand for housing. That is

$$p_j^s = \frac{1}{A_h}\left(\frac{\mu(j)}{\overline{H}_j}\right)^{\alpha_j}. \tag{1.39}$$

I then calculate the error, defined as the absolute maximum difference between the vector p^s and the initial price guess p. I terminate the program if this error is sufficiently small. If not I update the price guess by taking a weighted average between the two vectors as $p' = (1-\phi)p + \phi p^s$ and restart the entire iteration. [39] I continue iterating until I reach the preset error.

[38] While I could allow for education to vary with age, in this model death occurs at the same deterministic time for all households. So, for a steady-state I need the measure of households with education e to be constant over age.

[39] If the error grows over successive iterations ϕ is reduced, until we begin to converge.

1.B Solving for transitional dynamics

To solve for the transitional dynamics between two steady-states we need to find a sequence of prices for points along a path which begins at the pre-shock steady-state price and ends at the post-shock steady-state prices. The initial distribution of households across cities and wealth is taken to be the initial steady-state values. I create the series for the amenities, productivity, etc. in each city at each time period during the transition. Even after all the shocks have happened we are not necessarily at the steady-state as the endogenous distribution of households over regions and wealth may continue to change. I therefore add many periods after the shocks, amounting to 95 years, during which there are no further shocks.

The algorithm for solving for the price path along the transition is as follows. I first create an initial price guess by interpolating between the two steady-state prices. I use equation 1.18 to create the initial guess for city rents. I begin at the first period and solve for equilibrium prices under the assumption that all future prices are correct. To do this I solve for the households problem using equations 1.11, which contains the expected value conditional on choices, bond holdings and future prices. The conditional expected value is computed by backwards induction beginning in the final period of life. So, for a household with a periods of life left we need to solve 1.11 a times, each time using the previous value.

After completing the household's problem at time zero of the shock I then calculate the conditional choice probabilities using 1.10. Using the initial steady-state measures of households in cities conditional on a, e, and the initial steady-state bond-wealth distribution I compute the measure of households in each location conditional on age and education as well as the average bond-wealth conditional on age, education and location. Using this measure I can calculate the unconditional measure of households in each location, conditional on the initial price guess.

If the difference between the price guess and the price implied by 1.19 is sufficiently small I move to the next period. If not I then update the price, but *only* for period zero, using a weighted average of the current assumed price for period zero and the price implied by 1.19 for period zero. When resolving note that only period zero and period one see changes in rents or house prices, reducing computational needs.

Once period zero has converged to a preset tolerance I continue to period one. I solve for the households problem under the, now using the equilibrium prices and distributions from our solution in period zero. Solving the households problem is similar to period zero. However, I have already computed many of the conditional expected future values, which reduces the computation burden. I proceed in computing all future periods in a similar fashion.

After reaching the end of the time periods in the transition I calculate the absolute maximum difference between the initial price guess and the updated price guess across all cities and all time periods. If this difference is sufficiently small we have converged. Otherwise we restart the inner loop at time zero with the new updated price as the initial price guess. We continue this process until either the difference in prices between subsequent iteration is sufficiently small or we exceed a predetermined maximum number of iterations.

1.C Sorting Results

Here I outline a simplified version of the model to illustrate the mechanisms that generate the positive assortative matching between worker education and high-cost high-productivity regions. That is that more highly educated workers locate in these high-cost high-productivity regions with a higher probability. high-productivity workers are more likely to choose to locate in high as surplus from a worker locating in a region is increasing in education and regional productivity.

To see this formally, note that from equations (1.8) and (1.10) we can write that the probability

$$\pi(j'|a,b,e,j;S) = \frac{\exp(\log(c_{j'}) + \omega\chi_{j'} - \tau_{jj'} + \beta\mathbb{E}[v(a+1,b',e,j')])^{\frac{1}{\theta}}}{\sum_l \exp(\log(c_l) + \omega\chi_l - \tau_{jl} + \beta\mathbb{E}[v(a+1,b_l,e,l))^{\frac{1}{\theta}}} \tag{1.40}$$

To simplify matters I examine renters with a binding borrowing constraint so that $b = 0$. Then for renters consumption is $c = A_{j'}e - p_{rj'}$ and for homeowners $c = A_{j'}e - p_{j'} + (1-\delta)p_j - \psi p_j + \frac{\psi}{1+r}p_{j'}$. To see the assortative matching let $\frac{p_k}{p_l} > \frac{A_k}{A_l} > 1$. Then for renters we have that

$$\frac{\frac{\pi(a,0,e',j,k;S)}{\pi(a,0,e',j,l;S)}}{\frac{\pi(a,0,e,j,k;S)}{\pi(a,0,e,j,l;S)}} = \left(\frac{\frac{(A_k f(a',e') - p_{rk})\exp(\mathbb{E}[v(a,0,e',k)])}{(A_l f(a',e') - p_{rl})\exp(\mathbb{E}[v(a,0,e',l)])}}{\frac{(A_k f(a,e) - p_{rk})\exp(\mathbb{E}[v(a,0,e,k)])}{(A_l f(a,e) - p_{rl})\exp(\mathbb{E}[v(a,0,e,l)])}}\right)^{\frac{1}{\theta}} \tag{1.41}$$

In the final period we the expected value term is zero. This simplifies this ratio to

$$\left(\frac{\frac{(A_k f(a',e') - p_{rk})}{(A_l f(a',e') - p_{rl})}}{\frac{(A_k f(a,e) - p_{rk})}{(A_l f(a,e) - p_{rl})}}\right)^{\frac{1}{\theta}} \tag{1.42}$$

Differentiating this with respect to $f(a', e')$ one can show that

$$\frac{\partial}{\partial f(a',e')}\frac{\frac{\pi(a,0,e',j,k;S)}{\pi(a,0,e',j,l;S)}}{\frac{\pi(a,0,e,j,k;S)}{\pi(a,0,e,j,l;S)}} = \frac{1}{\theta}\left(\frac{A_l f(a,e) - p_{rl}}{A_k f(a,e) - p_{rk}}\right)^{1/\theta}\left(\frac{A_k p_{rl} - A_l p_{rk}}{(A_l f(a',e') - p_{rl})^2}\right)^{\frac{1-\theta}{\theta}} > 0 \tag{1.43}$$

Since this derivative is increasing in $f(a', e')$ workers with a higher effective labor sort into regions with higher productivity. A similar result can show that this holds for homeowners as well.

The crucial economic force that drives sorting in the model is the non-homothetic demand for housing services. In the model, any household that lives in region j demands one and only one home in the region. Thus, within a region the proportion of expenditure on housing falls as wealth increases. This leads to sorting of more educated households into higher productivity, high housing cost regions. This because when the sorting condition holds, i.e. $\frac{p_k}{p_l} > \frac{A_k}{A_l} > 1$, and $f(a', e') > f(a, e)$ then

$$\frac{A_k f(a',e') - p_k}{A_j f(a',e') - p_j} > \frac{A_k f(a,e) - p_k}{A_j f(a,e) - p_j} \tag{1.44}$$

So that the ratio of after housing costs income between the two regions is larger for the highly educated worker. Which of course implies that the more highly educated workers are more likely to locate in these high-productivity regions.

CHAPTER 2

Corporate Bond Liquidity during the COVID-19 Crisis

This paper was jointly written with Mahyar Kargar, Benjamin Lester, Shuo Liu, Pierre-Olivier Weill, Diego Zúñiga and published as Mahyar Kargar, Benjamin Lester, David Lindsay, Shuo Liu, Pierre-Olivier Weill, Diego Zúñiga, Corporate Bond Liquidity during the COVID-19 Crisis, The Review of Financial Studies, Volume 34, Issue 11, November 2021, Pages 5352–5401.

2.1 Introduction

The COVID-19 pandemic induced an unprecedented shock to the global economy. As the implications of this shock began to crystalize in mid-March, 2020, financial markets plummeted and reports of illiquidity began to surface. One particularly important market that was "under significant stress" (Bernanke and Yellen, 2020) was the $10 trillion corporate bond market, which a March 18 report from Bank of America deemed "basically broken" (Idzelis, 2020). In response, the Federal Reserve introduced several facilities designed to bolster liquidity and reduce the costs and risks of intermediating corporate debt, including the Primary Dealer Credit Facility (PDCF) and the Primary and Secondary Market Corporate Credit Facilities (PM-

CCF and SMCCF, respectively). The latter two facilities represented a particularly bold intervention, in that they allowed the Fed, for the first time, to make outright purchases of investment-grade corporate bonds issued by US companies, along with exchange-traded funds (ETFs) that invested in similar assets.

The purpose of this paper is to study trading conditions in the US corporate bond market in response to the large economic shock induced by COVID-19, as well as the unprecedented interventions that followed. Given the exogenous nature of the pandemic, set against the backdrop of a well-capitalized financial sector, this episode offers a unique opportunity to identify the nature of shocks that precipitate illiquidity in financial markets, the consequences for market participants, and the efficacy of various policy responses designed to restore liquidity in times of distress.

A central feature of our analysis is the distinction between two types of transactions offered by dealers: "risky-principal" trades, in which a dealer offers a customer-seller immediacy by purchasing the asset directly and storing it on his balance sheet until finding a customer-buyer; and "agency" trades, in which the customer-seller retains the asset while waiting for a dealer to find a customer-buyer to take the other side of the trade. This distinction, which has been studied recently using pre-pandemic data, is crucial in generating several new insights.[1] We highlight three.

First, distinguishing between risky-principal and agency trades provides a more complete assessment of market liquidity by accounting for both the cost of trading and the time it takes to trade. Indeed, we show that focusing on transaction costs

[1]For recent work that studies the distinction between risky-principal and agency trades, see Schultz (2017), Bao et al. (2018) , Choi and Huh (2018), Bessembinder et al. (2018), and Goldstein and Hotchkiss (2020a). To the best of our knowledge, we are the first (and only) paper to employ this distinction to study liquidity conditions during the COVID-19 pandemic.

alone—ignoring changes in the composition of risky-principal and agency trades—understates the deterioration in liquidity after the COVID-19 shock. Second, studying the cost and quantity of these distinct types of trades in concert with a structural model allows us to disentangle two widely cited (but not mutually exclusive) sources of illiquidity: a large, unexpected increase in customers' demand for immediacy, sometimes called a "dash for cash"; and a decrease in dealers' willingness to supply immediacy by absorbing assets onto their balance sheets, either because of rising costs or binding regulatory constraints. A key finding is that matching the data requires large shocks to *both* demand and supply at the onset of the crisis. Finally, studying the evolution of demand and supply factors against the timeline of the Fed's interventions offers new insights into the efficacy of various policies. In particular, we show that the surge in customers' demand for immediacy receded almost immediately, and fully, after the announcement of the Fed's interventions, whereas the negative shock to dealers' supply of immediacy responded more gradually, and only partially.

After providing some background information in Section 2.2, we begin our analysis in Section 2.3 by documenting trading conditions in the corporate bond market in response to the panic of mid-March and the Fed's interventions that followed. Using data from the Trade Reporting Compliance Engine (TRACE), we first construct time series to measure the costs of risky-principal and agency trades in the corporate bond market. We find that the cost of risky-principal trades increased significantly during the COVID-induced panic, reaching a peak of more than 250 basis points (bps), while the cost of agency trades increased much more modestly. As the premium paid for risky-principal trades increased, we show that customers substituted towards agency trades: the fraction of total volume executed as agency trades increased by as much as 15% at the height of the sell-off, and remained elevated even months after the initial panic subsided. Hence, the average trade was not only more expensive, but also more likely to be slower or of "lower quality."

As trading shifted from risky-principal to agency transactions, we show that the dealer sector as a whole absorbed *no* inventory, on net, during the most tumultuous period of trading. Therefore, when the demand for transaction services surged, it was customers themselves who ultimately stepped up to provide additional liquidity. In fact, it was only after the announcement of the Federal Reserve's interventions that dealers began to absorb inventory onto their balance sheets, and trading conditions started to improve. Indeed, after the announcement of the Fed's credit facilities, the quantity of corporate debt held by dealers more than doubled relative to pre-COVID levels. At the same time, the cost of risky-principal trades decreased significantly, but remained approximately twice the levels observed before the pandemic.

While these observations establish the coincidence of key interventions and improvements in market liquidity, they do not establish a causal relationship. To further explore the effects of interventions on market liquidity, we exploit restrictions on the types of bonds that could be purchased through the Fed's corporate credit facilities. In particular, using a difference-in-differences approach, we use restrictions on bond ratings and time-to-maturity to identify the change in trading costs induced by the announcement of the SMCCF. We find that, immediately after the announcement of the SMCCF, the cost of trading bonds that were eligible for purchase by the Fed decreased substantially relative to the cost of trading ineligible bonds. Later, when the program was expanded in both size and scope, we show that the trading costs of all bonds fell.

Hence, our findings suggest that the Fed's interventions had significant effects on transaction costs and trading activity in the corporate bond market. However, the observations described above also lead to several important questions. Why did the Fed's interventions improve trading conditions so quickly, but not fully? Did the announcement and implementation of these policies (at least partially) restore liquidity by easing investors' concerns and halting the "dash for cash"? Or should the efficacy of these interventions be attributed to easing dealers' balance sheet concerns, thus increasing their willingness to "lean against the wind" (Weill, 2007)? Given the deterioration of both the cost *and quality* of intermediation services during the COVID-19 crisis, what was the effect on the surplus of customers in the US corporate bond market during this period?

To confront these questions, and interpret our empirical findings more generally, in Section 2.4 we construct a parsimonious equilibrium model of a market for vertically differentiated transaction services: low-quality, meant to capture agency trades; and high-quality, meant to capture risky-principal trades. We assume that customers prefer high-quality transaction services, but they are more costly for dealers to produce. Within this framework, we characterize the differential impact of two types of shocks—to customers' relative demand for high-quality risky-principal trades and to dealers' cost of supplying these transaction services—on equilibrium prices and allocations.

Then, using our estimates of relative prices and quantities in concert with our theoretical framework, we estimate key parameters of the model, which allows us to identify shocks to customers' demand for immediacy, at the height of the crisis and during the interventions that followed. We confirm that a large, sudden increase in the demand for immediacy was a crucial source of illiquidity early in the crisis. In fact, we estimate that customers' willingness to pay for each inframarginal unit of risky-principal trade (rather than an agency trade) increased by approximately 200 bps at the height of the crisis. However, we also find that this shock alone cannot explain what we observe in the data: to rationalize the observation that customers ultimately substituted towards agency trades, we show that there must have also been a significant shock to dealers' marginal cost of supplying immediacy. Hence, understanding the market turmoil of March 2020 requires studying *both* the origins of the "dash for cash" *and* the factors that dissuaded dealers from absorbing selling pressure onto their balance sheets.

Studying the behavior of shocks to customers' demand for immediacy, and dealers' cost of supplying it, against the timeline of policy announcements and implementation also reveals new insights regarding the channels through which the Fed's interventions operated. In particular, we find that the demand shock receded quickly, and fully, soon after the *announcement* of the PDCF and SMCCF—that is, the announcement alone seems to have effectively reversed the initial "dash for cash." The increase in dealers' cost of supplying risky-principal trades, however, appears to have lingered even after the Fed began purchasing bonds. While there are multiple explanations for this, we document one plausible candidate: the total volume of customer-dealer transactions remained elevated through June which—when combined with binding balance sheets constraints—could explain why the relative cost (fraction) of risky-principal trades remained elevated (depressed) months after markets appear to have calmed.

Finally, we leverage our theoretical framework—along with our empirical estimates of preference shocks, prices, and quantities—to construct a measure of customers' well-being. In particular, we define consumers' *surplus from immediacy* as the net utility per unit of transaction that a customer receives from upgrading from slower, agency trades to faster, risky-principal trades. Relative to the pre-crisis period, we find that the loss in consumers' surplus from immediacy was less pronounced than the increase in the relative price premium for immediacy, but remained suppressed well after markets had calmed. In fact, we find customers' net utility from upgrading to faster, risky-principal trades per unit of transaction declined by only about 20 bps during the height of the market turmoil, but remained approximately 10 bps below pre-crisis levels even at the end of June, 2020. We argue that these results highlight the importance of accounting for changes in customers' preferences for immediate trades, along with changes in the relative quantities of risky-principal and agency trades, when assessing the effects of shocks and the interventions that follow.

Related literature

Given the size of the COVID-19 shock, and the historic nature of the Fed's response, it is not surprising that a number of recent papers have emerged to study financial markets since the onset of the pandemic. Our paper belongs to the literature focused on the corporate bond market, which we discuss in more detail below, but shares much in common with studies of other markets, including the market for Treasuries and other government debt (Duffie, 2020; He et al., 2020; Fleming and Ruela, 2020; Schrimpf et al., 2020), as well as the market for asset-backed securities (Foley-Fisher et al., 2020; Chen et al., 2020).

In the corporate bond market, Falato et al. (2020) study the effect of the pandemic on outflows from bond mutual funds, and the role that the Fed's corporate credit facilities played in reversing these outflows. Ma et al. (2020) also explore outflows in fixed-income mutual funds, including those that invest in corporate bonds and Treasuries. They derive a pecking order theory of liquidation, which explains why selling pressure was strongest in the most liquid sectors of these markets. Haddad et al. (2020) focus primarily on the behavior of credit spreads during the crisis, and attempt to identify the mechanism through which the Fed's interventions improved market conditions.[2] Though different along many dimensions, these three papers all argue that a large, sudden increase in customers' demand for immediacy played a crucial role in the deterioration of market liquidity in March, 2020. We, too, identify such a shock, but find that matching the data also requires a significant shock to the dealers' cost of supplying immediacy.

Our paper is most closely related to contemporaneous work by O'Hara and Zhou (2020) and Boyarchenko et al. (2020), who also investigate liquidity conditions in the corporate bond market during the COVID-19 crisis, and the effects of the Fed's interventions.[3] Despite some overlap, the three papers differ (and complement one another) in several important ways. For example, using the regulatory version of TRACE—which contains dealer identities—O'Hara and Zhou (2020) document the heterogeneous response of different dealers to the Fed's interventions. This allows them to control for dealer fixed effects and to disentangle the effects of the PDCF and the SMCCF, among other things. Boyarchenko et al. (2020) also use the reg-

[2]For related work on the behavior of credit risk/spreads throughout the crisis, and the effects of the Fed's interventions, see Nozawa and Qiu (2020) and D'Amico et al. (2020).

[3]In more recent work, Gilchrist et al. (2020) quantify the effects of the SMCCF on credit spreads and transaction costs using a regression discontinuity approach. Using a different methodology to construct their sample, they find qualitatively similar results regarding the effects of the SMCCF on transaction costs for eligible and ineligible bonds, though their quantitative magnitudes differ from ours. We discuss this further in Section 2.3.

ulatory version of TRACE, along with data on the volume of bonds (or shares of ETFs) purchased by the Fed's corporate credit facilities. This allows them to decompose the effects of the Fed's interventions into direct "purchase effects" and indirect "announcement effects."[4]

While our paper makes a number of distinct contributions relative to these contemporaneous studies, we highlight several aspects of our analysis that are particularly important. First, our approach to measuring trading conditions accounts for two channels through which market liquidity can deteriorate—customers can face higher transaction costs or longer waiting times for executing a trade—and hence provides a multi-dimensional assessment of market conditions during the crisis. Second, in contrast with the papers cited above, we develop a theoretical framework that, when combined with our empirical estimates, allows us to construct quantitative estimates of the shocks that precipitated the COVID-19 crisis.[5] Finally, we use these estimates of shocks to demand, along with bounds on shocks to supply, to study the efficacy of various policy interventions, and the implications for consumer surplus, at the height of the crisis and beyond.

[4]For more on the purchase effects of the SMCCF, see Flanagan and Purnanandam (2020).

[5]Along this dimension, our paper is related to Goldberg and Nozawa (2020), who use a structural VAR approach to identify demand and supply shocks in the corporate bond market during (and after) the 2007-2009 financial crisis. However, our identification strategy is different from theirs, as is our focus: their primary concern is the asset pricing implications of liquidity supply shocks.

2.2 Background

2.2.1 The COVID-19 Shock

Despite reports of a potentially lethal virus spreading in China, US equity markets reached all-time highs on February 19, 2020. Just two weeks later, as the scope of the COVID-19 coronavirus and the duration of its effects became apparent, financial markets around the world entered a period of turmoil. For example, between March 5 and March 23, the S&P 500 fell more than 25%. In the corporate bond market, the ICE Bank of America AAA US Corporate Index Option-Adjusted spread increased by about 150 bps over this same period, while the corresponding spread for high-yield (HY) corporate debt increased by more than 500 bps.[6] As the price of equities and debt plummeted, reports of illiquidity in key financial markets emerged. Such reports were especially troubling in the corporate bond market, as many large US firms would almost surely need access to capital in light of the impending shocks to their balance sheets.[7]

Two complementary factors were cited as the root of the panic in the corporate bond market. The first was a surge in the demand for immediacy, or so-called "dash for cash," as investors pulled out of corporate bond funds in droves. For example, Falato et al. (2020) report that, between the months of February and March, the average corporate bond fund experienced cumulative outflows of approximately 9% of net asset value—by far the largest outflows in the last decade. At the same time,

[6]See Ebsim et al. (2020) for a more comprehensive analysis of credit spreads during this time period.

[7]Indeed, Darmouni and Siani (2020) document that corporate bond issuance reached historic levels in the Spring of 2020, after the Fed's interventions, despite a relatively healthy banking sector.

market participants reported that dealers were either unable or unwilling to supply customers with immediacy by absorbing corporate debt onto their balance sheet. In a *Wall Street Journal* article titled "The Day Coronavirus Nearly Broke the Financial Markets," Baer (2020) writes about the experience of Vikram Rao, the head bond trader of Capital Group, after calling senior executives for an explanation on why broker-dealers wouldn't trade:

> [T]hey had the same refrain: There was no room to buy bonds and other assets and still remain in compliance with tougher guidelines imposed by regulators after the previous financial crisis [...] One senior bank executive leveled with him: "We can't bid on anything that adds to the balance sheet right now."

2.2.2 Federal Reserve Interventions

In response to signs of illiquidity in several key financial markets, the Federal Reserve introduced a number of new facilities designed to bolster liquidity and reduce trading costs. On the evening of March 17, the Federal Reserve revived the aforementioned PDCF, offering collateralized overnight and term lending to primary dealers. By allowing dealers to borrow against a variety of assets on their balance sheets, including investment-grade corporate debt, this facility intended to reduce the costs associated with holding inventory and intermediating transactions between customers.[8]

---

[8] In addition to the facilities that we highlight in our analysis here, it is also noteworthy that the Federal Reserve temporarily relaxed the supplementary leverage ratio (SLR) rule—first on April 1 and again on May 15, 2020—to ease balance sheet constraints and increase banks' ability to lend to households and businesses. By excluding US Treasury securities and reserves from

On March 23, the Federal Reserve proposed even more direct interventions in the corporate bond market through the PMCCF and SMCCF. These facilities were designed to make outright purchases of corporate bonds issued by investment-grade US companies with remaining maturity of five years or less. The facilities were also allowed to purchase shares in US-listed exchange-traded funds (ETFs) that invested in US investment-grade corporate bonds. On April 9, these corporate credit facilities were expanded in size and extended to allow for purchases of ETFs that invested in high-yield corporate bonds.[9] Interestingly, though many of the effects of these corporate credit facilities were observed immediately after they were announced (and expanded), the Federal Reserve did not actually begin purchasing bonds until May 12. We provide a more detailed description of this timeline, and of the Federal Reserve's corporate facilities, in Appendix 2.B.

2.3 Trading Conditions During the Pandemic

In this section, we describe how market conditions evolved from the sanguine conditions of mid-February through the freefall of mid-March to the post-intervention recovery of April and May. As a first step, we construct time series for several variables of interest: the cost of risky-principal trades, the cost of agency trades, and

---

the calculation of the SLR rule for holding companies, the rule change was primarily intended to increase liquidity in the Treasury market. However, to the extent that it relaxed dealers' balance sheet constraints, the effects could clearly extend to the corporate bond market as well, as we discuss later in the text. To read more about the rule change, see press releases on April 1, 2020 (https://www.federalreserve.gov/newsevents/pressreleases/bcreg20200401a.htm) and May 15, 2020 (https://www.federalreserve.gov/newsevents/pressreleases/bcreg20200515a.htm).

[9] The April 9 update also allowed the SMCCF to make direct purchases of bonds that had been downgraded from investment-grade to high-yield status (so-called "fallen angels") after March 22. The facility also allowed purchasing of high-yield ETFs.

the fraction of each type of transaction services. We document that, at the height of the selling pressure, the cost of risky-principal trades surged and the fraction of such trades dropped significantly. Conditions improved immediately after the Fed's announcement of the corporate credit facilities, with dealers providing liquidity directly, via risky-principal trades, at significantly lower prices. To test the causal relationship between the Fed's interventions and market liquidity, we exploit the eligibility requirements for bond purchases by the SMCCF. We find that, after the initial announcement, trading costs for eligible bonds fell substantially more than trading costs for ineligible bonds. Later, after the program was expanded in both size and scope, we document more significant declines in trading costs for all bonds.

2.3.1 Data

We combine the standard TRACE data set (for 2020Q1) with the End-of-Day version (for 2020Q2). We first filter the report data following the standard procedure laid out in Dick-Nielsen (2014). We merge the resulting data set with the TRACE master file, which contains bond grade information, and with the Mergent Fixed Income Securities Database (FISD) to obtain bond fundamental characteristics. Following the bulk of the academic literature, we exclude variable-coupon, convertible, exchangeable, and puttable bonds, as well as asset-backed securities, and private placed instruments. We also exclude newly-issued and foreign securities.

For most of our analysis, we use the (filtered) data covering the period from January 2 to June 30, 2020, which contains 7.2 million trades and 30,748 unique bonds. Approximately 60% of the transactions are identified as customer-dealer and 40% as interdealer trades. The average trade size is $218,104 across all transactions, with average total daily volumes for customer-dealer and interdealer trades of \$7.25 billion and \$3.13 billion, respectively. It is worth noting that, in both the standard and End-of-Day versions of TRACE, the trade size for investment-grade and high-yield bonds is top-coded at \$5 million and \$1 million, respectively.[10]

In all of our plots below, we include vertical dashed lines to highlight several key dates: February 19, when stock markets reached their all-time peaks; March 5, which marks the beginning of the extended fall in equity prices and rise in corporate credit spreads; March 18, the first day of trading after the announcement of the PDCF; March 23, the day that the PMCCF and SMCCF were announced; April 9, the day that the size and scope of the corporate credit facilities were expanded; May 12, the date that the SMCCF started buying bond ETFs; June 16, the day that the SMCCF began purchasing individual bonds; and June 29, the date the PMCCF began operating.

2.3.2 The cost of trading, fast and slow

To capture the average transaction cost for risky-principal trades, we use the measure of bid-ask spreads proposed by Choi and Huh (2018), CH hereafter. To construct this measure, we first calculate, for each customer trade, the spread

$$2Q \times \frac{\text{traded price} - \text{reference price}}{\text{reference price}}, \tag{2.1}$$

where Q is equal to $+1$ (-1) when a customer buys from (sells to) a dealer, and the reference price is taken to be the volume-weighted average price of interdealer

[10] Table 2.A.1 in Appendix 2.A presents additional summary statistics for our sample.

trades larger than \$100,000 in the same bond-day. Importantly, we restrict our sample so that it only includes trades in which the dealer who buys the bond from a customer holds it for more than 15 minutes. In doing so, we leave out those trades where the dealer had pre-arranged for another party (either a customer or another dealer) to buy the bond immediately.[11] The measure of risky-principal trading costs is aggregated at the bond-day level by taking the volume-weighted average of trade level spreads, and then at the daily level by taking the average in each day across all bonds, weighted by bonds' daily total volume of customer trades where the CH measure is available.

To capture the average transaction cost of agency trades, we calculate a modified version of the Imputed Roundtrip Cost measure described in Feldhütter (2012). To construct this modified imputed roundtrip cost (or "MIRC"), we first identify imputed roundtrip trades (IRT) by matching a customer-sell trade with a customer-buy trade of the same size that takes place within 15 minutes of each other.[12] We exclude interdealer trades in constructing IRTs, so that each IRT only includes one customer-buy trade and one customer-sell trade. Then, to compute the MIRC, we calculate

$$\frac{P_{max} - P_{min}}{P_{max}},$$

where P_{max} (P_{min}) is the largest (smallest) price in the IRT. Within each bond and day, we calculate the daily average roundtrip cost as the average of the bond's MIRC

[11]Likewise, in calculating reference prices, we follow CH and exclude interdealer trades executed within 15 minutes of a customer-dealer trade.

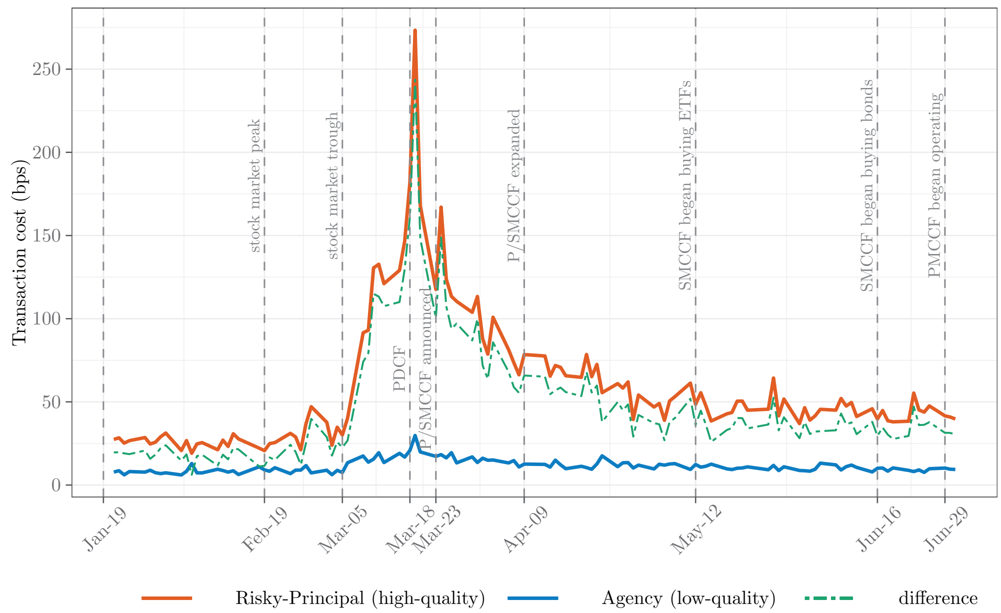
[12]In other words, as in earlier papers, we assume that customer-buys and customer-sells that occur in rapid succession are likely to be agency trades. Indeed, in an agency trade, dealers search for counterparties on behalf of customers. When counterparties are found, the two customers are matched by dealers, and two customer-to-dealer trades are recorded in a short time window.

on that day, weighted by trade size. Finally, a daily estimate of average roundtrip cost is the average of roundtrip costs on that day across all bonds, weighted by bonds' total daily trading volumes in the matched IRTs.

Figure 2.3.1 plots the two time series, along with the difference between the two. The two measures of transaction costs are relatively stable through February 19, with risky-principal trades approximately twice as expensive as agency trades. Upon realization of the COVID-induced shock, the cost of risky-principal trades rises dramatically, while the cost of agency trades is more muted. In particular, between Thursday, March 5, and Monday, March 9, the cost of risky-principal trades roughly triples, to approximately 100 bps; over these three trading days, the S&P 500 Index declined more than 12%. A week later, during the most tumultuous period of March 16-18, this series continues to rise, reaching a peak of more than 250 bps, before beginning a steady decline after the announcement of the SMCCF on March 23. The MIRC measure of agency trading costs, in contrast, increases from a baseline around 8 bps to approximately 28 bps, before receding slightly after the Fed's intervention.

One can see that the cost of risky-principal trades, which we interpret as the cost of trading immediately, was considerably more responsive to both the heightened selling pressure induced by the pandemic in mid-March and the Fed's interventions which followed. Moreover, despite considerable improvement in both metrics during the month of April, the cost of risky-principal trades remained elevated through June, which suggests that liquidity conditions remained somewhat strained well after markets appear to have calmed.

Of course, the change in spreads could be driven by a change in the composition of bonds that were traded during this period of distress. For example, perhaps trading



2.3.1. Transaction costs: Risky-principal vs. agency trades. This figure shows the time-series of trading costs for risky-principal trades in red and agency trades in blue, and their difference in green.

volume was unusually high for retail-size trades of illiquid bonds, which typically involve higher transaction costs. Thus, to further clarify the impact of the crisis and ensuing interventions on the cost of risky-principal and agency trades, we turn to formal regressions that allow us to control for bond- and trade-level fixed effects. We consider the following specification

$$y_{ijt} = \alpha_i + \alpha_s + \beta_1 \times \text{Crisis}_t + \beta_2 \times \text{Intervention}_t + \varepsilon_{ijt}. \tag{2.2}$$

The dependent variable, y_{ijt}, represents the transaction cost for a type $j \in \{\text{risky-principal, agency}\}$ trade of bond i on day t. The dummy variables Crisis_t and Intervention_t allow us to distinguish between three sub-periods: (i) Pre-crisis, which corresponds to dates before March 5, 2020; (ii) Crisis, which covers the period March 5–23, 2020; and (iii) Intervention, which covers the period after March 23. Hence, the coefficients β_1 and β_2 measure transactions costs relative to the pre-crisis period. Finally, α_i and α_s represent bond and trade size fixed effects, respectively. Bond fixed effects capture bond characteristics that are fixed over time such as industry, par amount, etc.[13] For trade size fixed effects we consider three categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million.

Table 2.3.1 presents results for all bonds, as well as the sub-sample of bonds issued by US firms.[14] We include bond and size category fixed effects and cluster standard errors at the bond and day levels in all regressions to account for correlation over time within a bond and across bonds in a given day. Columns (1) and (3) reveal that, during the crisis period of March 5-23, average bond-level trading costs for risky-principal and agency trades increased by approximately 105 bps and 9 bps, respectively, relative to the pre-crisis period. After the Fed's interventions on March 23, trading costs for risky-principal trades fell by approximately 64 bps—more than half the initial spike—while transaction costs for agency trades declined much more modestly. These results are consistent with the aggregate results in Figure 2.3.1. Columns (2) and (4) show that the sub-sample of US-issued bonds exhibits roughly the same behavior as the sample of all bonds, though the cost of agency trades for US-issued bonds increased slightly more during the crisis period.

---

[13]We do not have access to the latest credit rating data for all bonds in our sample, just the binary IG/HY classification provided by TRACE. For the sub-sample of bonds where the credit rating is available, we include a credit rating fixed effect in specification (2.2) to control for potentially time-invariant nature of bond credit ratings.

[14]One reason we include the results for the US sub-sample is to demonstrate that the trading cost patterns are similar to the full sample. This is helpful later, in Section 2.3.5, when we focus on the US sub-sample exclusively.

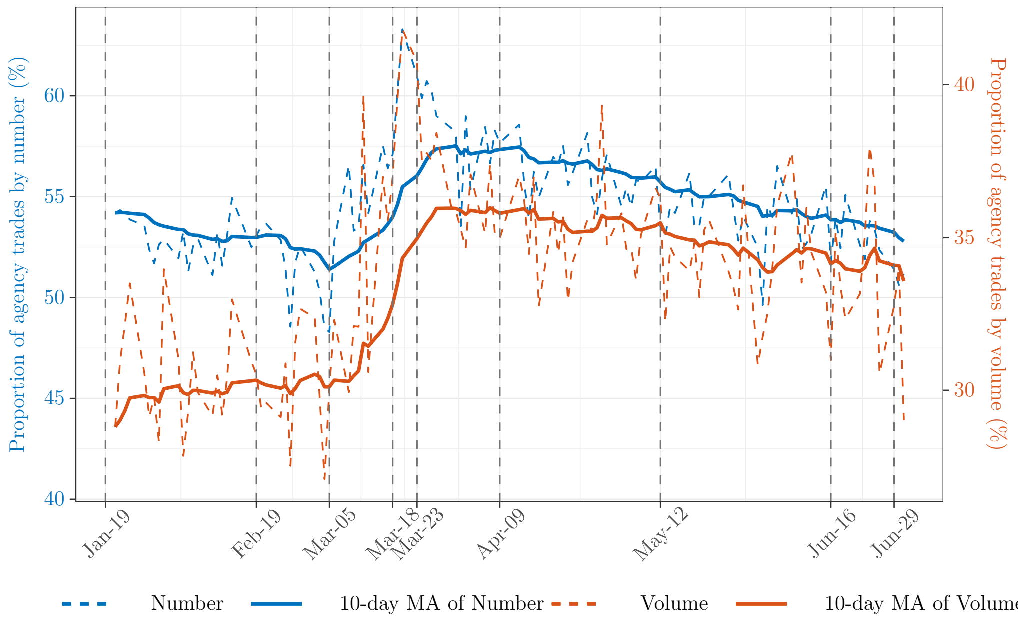
2.3.1. Trading costs during the COVID-19 crisis. This table presents regression results for the following specification: $y_{ijt} = \alpha_i + \alpha_s + \beta_1 \times \text{Crisis}_t + \beta_2 \times \text{Intervention}_t + \varepsilon_{ijt}$. The dependent variables are our measures of transactions costs for risky-principal and agency trades. Crisis_t and Intervention_t are dummies which take the value of 1 if day t falls into the Crisis and Intervention sub-periods defined above. There are three trade size categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million. The sample starts on January 3 and ends on June 30, 2020. Clustered standard errors at the day and bond levels are shown in parentheses. $^*p < .1$; $^{**}p < .05$; $^{***}p < .01$.

	Dependent variable:			
	Risky-principal		Agency	
	All	US only	All	US only
	(1)	(2)	(3)	(4)
Crisis	105.19***	104.76***	8.70***	9.99***
	(13.08)	(13.78)	(1.71)	(2.06)
Intervention	41.54***	40.34***	8.04***	8.26***
	(4.06)	(4.32)	(0.76)	(1.06)
Bond FE	Yes	Yes	Yes	Yes
Trade size category FE	Yes	Yes	Yes	Yes
Observations	741,579	603,913	249,392	160,539
Adjusted R^2	.18	.19	.28	.28

2.3.3 Substituting agency trades for risky-principal trades

We now establish that, as the premium for risky-principal trades increased, customers responded by substituting towards agency trades. Figure 2.3.2 plots the proportion of agency trades by number (left axis) and volume (right axis).[15] During the most tumultuous weeks of trading, between March 5 and March 23, the fraction of agency trades (measured by both number and volume) increased by as much as 15 percentage points, trough to peak, before receding after the March 23 announcement of the corporate credit facilities. Again, this shift toward agency trades has impor-

[15]We discuss how we identify agency trades in depth in Appendix 2.A.



2.3.2. Proportion of agency trades. This figure plots the fraction of agency trades by volume in red (right axis) and by number in blue (left axis).

tant implications for assessing market liquidity. In particular, if one were simply to measure trading costs across all trades, they would underestimate the erosion in liquidity as the composition of trades shifted from faster, more expensive risky-principal trades to less costly, but slower agency trades.

To study the substitution from risky-principal to agency trades more carefully, we consider a regression with the following specification:

$$\text{Agency}_{ijt} = \alpha_i + \alpha_s + \beta_1 \times \text{Crisis}_t + \beta_2 \times \text{Intervention}_t + \varepsilon_{ijt}, \tag{2.3}$$

where Agency_{ijt} is an indicator variable that takes the value one if trade j for bond i on day t is an agency trade and zero otherwise. The variables on the right-hand side of specification (2.3) are the same as in (2.2). Under this specification, the coefficients

β_1 and β_2 measure the change in the probability of an agency trade during the crisis and intervention periods, respectively, relative to the pre-crisis period. Table 2.3.2 presents results using a linear probability model (OLS), along with logit and probit specifications for robustness.

Column (1) reveals that, during the crisis period of March 5–23, the probability of an agency trade for a given bond, on average, rose by 4.3 percentage points relative to the pre-crisis period. After the Fed interventions on March 23, this probability decreased from the crisis period (by 160 bps) to 2.7 percentage points higher than the pre-crisis period. For the sake of completeness, we report marginal effects calculated at the sample means for logit and probit models in columns (2) and (3); the results are very similar to the linear probability model (OLS) in column (1).[16]

2.3.4 Dealers' inventory accumulation

As the relative price of risky-principal trades spiked in mid-March, and customers substituted towards agency trades, one might naturally wonder: who was providing liquidity in the corporate bond market? Were dealers "leaning against the wind" and absorbing some of the inventory during the selloff? Or was the shift to agency trades sufficiently large that other customers were ultimately providing liquidity? To answer this question, we construct a measure of the (cumulative) value of bonds that were absorbed over time by the dealer sector. In particular, using the daily Market Sentiment data from FINRA, we subtract the value of bonds that dealers sell to

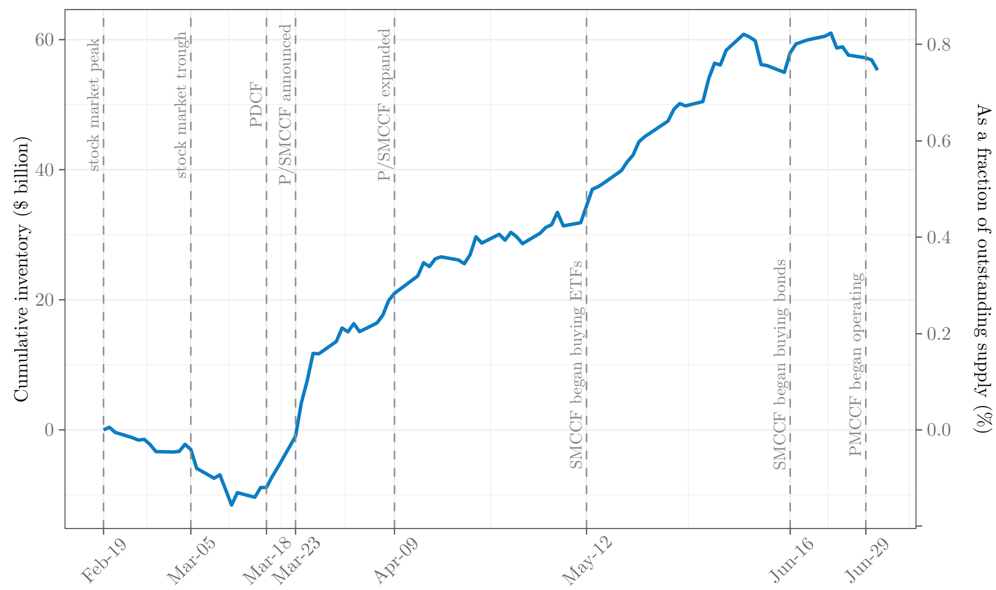
[16]For the interested reader, we also report results from a linear probability model that distinguishes between eligible and ineligible bonds for the SMCCF in Appendix 2.C. We find that the shift towards agency trades was more pronounced among bonds that were eligible for the Fed's purchasing program.

2.3.2. Probability of an agency trade for all bonds. This table presents regression results for the following specification from: $\text{Agency}_{ijt} = \alpha_i + \alpha_s + \beta_1 \times \text{Crisis}_t + \beta_2 \times \text{Intervention}_t + \varepsilon_{ijt}$. The dependent variable, Agency_{ijt}, is an indicator variable that takes the value 1 if trade j for bond i on day t is an agency trade and 0 otherwise. Columns (1), (2), and (3) report result for the linear probability (OLS), logit, and probit models, respectively. We report marginals effects calculated at the sample means for logit and probit models in columns (2) and (3). Crisis_t and Intervention_t are dummies which take the value of 1 if day t falls into Crisis and Intervention sub-periods defined above. There are three trade size categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million. In logit and probit specifications, the pseudo-R^2 is defined as $1 - L_1/L_0$, where L_0 is the log likelihood for the constant-only model and L_1 is the log likelihood for the full model with constant and predictors. The sample starts on January 3 and ends on June 30, 2020. Clustered standard errors at the day and bond levels are shown in parentheses. *p <.1; **p <.05; ***p <.01.

	Dependent variable:		
	Probability of agency trade		
	OLS	Logit	Probit
	(1)	(2)	(3)
Crisis	0.043***	0.043***	0.042***
	(0.010)	(0.010)	(0.010)
Intervention	0.027***	0.027***	0.027***
	(0.003)	(0.003)	(0.003)
Bond FE	Yes	Yes	Yes
Trade size category FE	Yes	Yes	Yes
Observations	7,095,617	7,095,617	7,095,617
Adjusted R^2	.104		
Pseudo-R^2		.079	.079

customers from the value of bonds that they buy from customers each day, and then calculate the cumulative sum of the net changes.[17] Figure 2.3.3 plots the cumulative

[17]The Market Sentiment data is available through FINRA TRACE Market Aggregate Information from https://finra-markets.morningstar.com/BondCenter/TRACEMarketAggregateStats.jsp. We use this data, as opposed to the standard or End-of-Day TRACE data, because it is not top-coded and hence allows for a more accurate assessment of the inflow and outflow of bonds in the dealer sector.



2.3.3. Cumulative inventory change in the dealer sector. This figure plots the cumulative inventory change in the dealer sector in billions of USD (left axis) and as a fraction of total supply in % (right axis). Source: FINRA market sentiment tables.

net change in inventory held in the dealer sector, both in levels (left axis) and as a fraction of pre-crisis outstanding supply (right axis).

Several aspects of Figure 2.3.3 are striking. First, during the most tumultuous period of trading, the dealer sector absorbed, on net, *no* additional inventory despite the considerable selling pressure from customers. In fact, dealers actually *reduced* inventory holdings and became net sellers. Hence, during this period, it was indeed other customers that were supplying liquidity to the market. Second, dealers' reluctance to absorb inventory appears to have changed substantially around the dates corresponding to the Fed's announcement of the Primary Dealer Credit Facility (March 18) and the Primary and Secondary Market Corporate Credit Facilities (March 23). Lastly, dealers continued to accumulate inventory through April and

May. Indeed, from March 18, the data indicates that dealers absorbed more than $50 billion in corporate debt, or roughly doubled their inventory holdings relative to pre-pandemic levels.[18]

2.3.5 The effects of the Fed's intervention

The results above suggest that the Fed's interventions—in particular, the March 23 announcement of the SMCCF—had a significant effect on dealers' willingness to absorb inventory onto their balance sheets, and hence on market liquidity. In this section, we exploit the eligibility requirements specified in the SMCCF to test this hypothesis more formally.

According to the original term sheet, a bond is eligible to be purchased through the SMCCF if it has an investment-grade rating on March 23, 2020; if it has a time-to-maturity of five years or less; and if its issuer is domiciled in the US.[19] However, the Fed has a considerable degree of discretion to determine whether a foreign issuer is domiciled in the US. Indeed, in the Fed's SMCCF transaction-level disclosures, we found many cases in which the holding firm of the security is a non-US entity.[20] Given this lack of clarity, we chose to focus on US firms exclusively, and classify a

[18]From Table L.130 of the Flow of Funds, at the end of 2019Q4, security brokers and dealers held $54 billion in corporate and foreign bonds on the asset side of their balance sheets.

[19]The original March 23 term sheet can be found at https://www.federalreserve.gov/monetarypolicy/smccf.htm. Initially, there was an additional eligibility criterion for the SMCCF on March 23: eligible issuers excluded firms that were expected to receive direct financial assistance from the then-pending CARES Act. This criterion (and others) were later added to the SMCCF term sheet on April 9. See Appendix 2.B for more details.

[20]SMCCF transaction-level disclosures are available at https://www.federalreserve.gov/monetarypolicy/smccf.htm. We provide additional details of this issue, including examples, in Appendix 2.A.

bond as eligible based on credit rating and time-to-maturity alone.[21]

To start, we repeat the regression specified in (2.2) with two modifications. First, we separate the sample of bonds into those that were eligible for purchase through the SMCCF and those that were not. Second, we separate the intervention period into two sub-periods. The first sub-period, which we call the "SMCCF," covers from March 23-April 8, 2020. During this period, it appeared that only investment-grade bonds would be eligible for purchase. The second sub-period, which we call the "SMCCF expansion," starts on April 9, when the Fed announced that it was increasing the size of the program and expanding the set of eligible bonds to include high-yield debt.

Table 2.3.3 reports the results. Column (2) reveals that the initial decline in trading costs was largely driven by bonds that were eligible for the SMCCF: the price of risky-principal trades for ineligible bonds declined much more modestly immediately after the March 23 announcement, relative to the crisis period, while the price of agency trades for ineligible bonds actually increased during this time period. After the program was expanded on April 9, in both scope and size, the price of risky-principal trades for all bonds declined significantly.

To further explore the causal effect of the SMCCF on bond market liquidity during the crisis, we consider a difference-in-differences regression over a sub-sample of our data from March 6 to April 9, 2020. These dates are chosen to exclude the pre-crisis period, when spreads were very low, and the post-expansion period, when the set of bonds available for purchase through the SMCCF was widened to include

[21] Recall from Table 2.3.1 that transaction costs for US firms behaved very similarly to all bonds in our sample.

2.3.3. Trading costs across eligible and ineligible bonds during the initial and expanded interventions. This table presents regression results for the following specification: $y_{ijt} = \alpha_i + \alpha_s + \beta_1 \times \text{Crisis}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{SMCCF expansion}_t + \varepsilon_{ijt}$. The dependent variables are measures of transaction costs for risky-principal and agency trades. Crisis_t is a dummy which takes the value of 1 if day t falls into the Crisis sub-periods defined above. SMCCF_t and SMCCF expansion_t are dummies that take the value of 1 if the trading day t is between March 23 and April 9, and after April 9, 2020, respectively. The SMCCF eligibility criteria were expanded to include fallen angels on April 9, 2020. There are three trade size categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million. A bond is considered eligible if it has an investment-grade rating and time-to-maturity of five years or less on March 23, 2020. The sample begins on January 3 and ends on June 30, 2020. Only US firms are included in the regressions. Clustered standard errors at the day and bond levels are shown in parentheses. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

	Dependent variable:					
	Risky-principal			Agency		
	All	Eligible	Ineligible	All	Eligible	Ineligible
	(1)	(2)	(3)	(4)	(5)	(6)
Crisis	107.97***	108.88***	106.67***	10.55***	15.94***	7.39***
	(14.77)	(15.14)	(16.60)	(2.30)	(3.35)	(1.91)
SMCCF	83.17***	61.94***	96.63***	13.43***	11.97***	14.43***
	(7.99)	(8.37)	(9.82)	(0.96)	(1.22)	(1.36)
SMCCF expansion	26.55***	14.15***	33.05***	6.09***	4.25***	7.31***
	(2.69)	(2.60)	(3.63)	(0.88)	(0.92)	(1.21)
Bond FE	Yes	Yes	Yes	Yes	Yes	Yes
Trade size category FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	602,430	219,624	382,806	159,653	56,264	103,389
Adjusted R^2	.19	.18	.19	.29	.21	.29

high-yield bonds. In particular, we use the specification

$$y_{ijt} = \alpha_s + \alpha_k + \beta_1 \times \text{SMCCF}_t \times \text{Eligible}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{Eligible}_t + \gamma \times X_{i,t} + \varepsilon_{ijt}, \quad (2.4)$$

where, as before, y_{ijt} represents our measures of transaction costs; Eligible_t takes the value of 1 if the bond in trade j has an investment-grade rating and time-to-maturity of five years or less on March 23, 2020; SMCCF_t takes the value of 1 if the trade occurs between March 23 and April 9, 2020; and α_s controls for size fixed effects.[22]

Unlike specification (2.2), we do not include bond fixed effects in the baseline specification (2.4), but instead control for industry fixed effects (α_k) and bond-specific characteristics such as bond age, amount outstanding, and time-to-maturity ($X_{i,t}$). However, for robustness, we also include results allowing for bond fixed effects, as well as credit rating fixed effects. To ensure that treatment and control groups do not overlap, we remove all trades in bonds that were downgraded from IG to HY. Finally, we drop all foreign bonds and focus only on bonds issued by US firms.

Table 2.3.4 contains our results. As is standard in difference-in-differences regressions, β_1 is the primary coefficient of interest. The first key takeaway is that the SMCCF had a significant effect on the cost of risky-principal trades for eligible bonds relative to ineligible bonds. The quantitative magnitude of this effect is approximately 50 bps, and is robust to a variety of alternative specifications. For example, in column (2) we include a credit rating fixed effect, which allows us to control for differences in average transaction costs using finer definitions of credit rating than IG or HY (e.g., AAA, AA, and so on), but has relatively minor effects on β_1. In columns (3) and (4), we allow for bond-specific fixed effects, which increases the explanatory power of the regressions (i.e., R^2) but does not significantly change

[22] One potential complication in distinguishing between eligible and ineligible bonds based on maturity is that the criteria for eligibility are determined at the Fed's time of purchase. Therefore, for example, a bond that would be characterized as ineligible when the SMCCF was announced on March 23, 2020, might, in fact, be purchased by the Fed in November 2020 (since the program remained active until December 31, 2020).

the estimates of β_1.

The second noteworthy result is that, for risky-principal trades, β_2 is not statistically different from zero under any of our specifications. Hence, it appears that the *announcement* of the initial SMCCF did not have significant spillover effects on the cost of risky-principal trades for ineligible bonds. However, this does not rule out the potential for spillover effects from the actual *purchase* of eligible bonds, which began on May 12, 2020. In particular, by purchasing bonds and relaxing dealers' balance sheet constraints, the SMCCF could potentially increase dealers' willingness to purchase any bond. If this is true, then some of the post-expansion decline in the costs of risky-principle trades for ineligible bonds (reported in Table 2.3.3) could be attributed to spillover effects from the Fed's bond purchases.

Columns (5)–(8) indicate that the announcement of the SMCCF on March 23 also reduced the cost of agency trades for eligible bonds.[23] One possible explanation is that, by establishing itself as a buyer of last resort, the Federal Reserve reduced the risk to private investors from purchasing eligible corporate bonds. According to this logic, it is possible that the announcement of the SMCCF made it easier for dealers to locate customer-buyers, hence reducing the spreads they charged on agency trades for eligible bonds. Note that this mechanism could also explain why the cost of agency trades for ineligible bonds went up in the immediate aftermath of the SMCCF announcement: if budget-constrained customers substituted from ineligible to eligible bonds, it would become more difficult for dealers to locate customer-buyers for ineligible bonds, driving spreads up.

In Appendix 2.C, we provide several additional robustness checks for the results

[23]Note that, looking at the overall effect ($\beta_1+\beta_2+\beta_3$), column (6) indicates that, after controlling for credit rating, the cost of agency trades for eligible bonds decreased after SMCCF announcement.

discussed above. In particular, in Tables 2.C.2 and 2.C.3, we show that the impact of the SMCCF on the trading cost of eligible bonds is even more pronounced if we limit our sample to those bonds that are just above and below the eligibility thresholds for and credit rating, respectively. In addition, in Tables 2.C.4–2.C.6, we show that small and large trades are responsible for the entire liquidity improvement documented in Table 2.3.4: small trades (with par volume of $100,000 or less) become much more liquid after the SMCCF announcements, while large trades (with volume larger than $1 million) also exhibit a significant decline in trading costs. Odd-lot trades (with volume between $100,000 and $1 million), however, are essentially unaffected by the Fed's intervention.

2.4 A structural analysis

The empirical analysis above highlights that the US corporate bond market experienced a significant decline in liquidity at the onset of the COVID-19 crisis, which was partially reversed by the Fed's interventions. Though informative, the facts we document leave several key questions unanswered. What was the nature of the shocks that led to a lack of liquidity? Why did the policies that were implemented appear to restore liquidity relatively quickly, but only partially? And how did these shocks and the ensuing interventions affect the well-being of the customers in this market?

To confront these questions, we now construct a parsimonious equilibrium model of the market for immediacy and use it to conduct a structural analysis of our empirical observations. Our analysis reveals that, at the onset of the crisis, the market was hit by large shocks to both customers' demand for immediacy (the "dash

2.3.4. The Effects of Fed Intervention: difference-in-differences. This table presents regression results for the following difference-in-differences specification from Equation (2.4): $y_{ijt} = \alpha_s + \alpha_k + \beta_1 \times \text{SMCCF}_t \times \text{Eligible}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{Eligible}_t + \gamma \times X_{i,t} + \varepsilon_{ijt}$. The dependent variables are measures of transaction costs for risky-principal and agency trades. SMCCF_t is a dummy that takes the value of 1 if day t falls between March 23 and April 9, and 0 otherwise. Eligible_t takes the value of 1 if the bond has an investment-grade rating and time-to-maturity of five years or less on March 23, 2020. X_{it} controls for log(Amt outstanding), log(Age), and log(Time-to-maturity): logs of bond's amount outstanding, years since bond issuance, and years to maturity, respectively. There are three trade size categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million. The sample begins on March 6 and ends on April 9, 2020. Only US firms are included and bonds that change credit grade are excluded. Clustered standard errors at the day and bond levels are shown in parentheses. $^*p<.1$; $^{**}p<.05$; $^{***}p<.01$.

	Dependent variable:							
	Risky-principal				Agency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SMCCF × Eligible	-57.70^{***}	-41.72^{***}	-47.24^{***}	-41.45^{***}	-10.25^{***}	-12.85^{***}	-9.59^{**}	-9.85^{***}
	(11.80)	(12.27)	(10.21)	(10.34)	(2.99)	(3.11)	(3.44)	(3.47)
SMCCF	-1.89	-21.75	-14.30	-20.03	6.33^{***}	8.10^{***}	4.56^{**}	4.72^{**}
	(14.58)	(14.64)	(14.65)	(14.43)	(2.00)	(2.11)	(1.97)	(2.02)
Eligible	2.86	-14.81			0.37	9.93^{***}		
	(14.24)	(11.36)			(3.15)	(3.69)		
log(Amt outstanding)	-30.33^{***}	-31.88^{***}			-3.62^{***}	-1.87^{***}		
	(7.25)	(9.19)			(0.64)	(0.65)		
log(Time-to-maturity)	15.40^{***}	16.77^{***}			4.00^{***}	5.53^{***}		
	(4.96)	(4.99)			(0.85)	(1.26)		
log(Age)	27.61^{***}	28.84^{***}			4.93^{***}	5.24^{***}		
	(7.54)	(6.40)			(1.10)	(1.14)		
Trade size category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	No	No	Yes	Yes	No	No
Bond FE	No	No	Yes	Yes	No	No	Yes	Yes
Credit rating FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	158,647	146,143	158,649	146,143	47,628	45,324	47,630	45,324
Adjusted R^2	.04	.05	.20	.20	.08	.10	.25	.26

for cash") and dealers' willingness to supply it. After the announcement of the Fed's key policy interventions, the demand shock subsided relatively quickly, and fully, while the supply shock recovered more gradually, and only partially. Relative to the pre-crisis period, we find that the loss in consumers' surplus from immediacy was less pronounced than the increase in the relative price of immediacy, $p_{ht} - p_{lt}$, but remained suppressed well after markets had calmed. In fact, we find customers' net utility from upgrading to faster, risky-principal trades remained approximately 10 basis points below pre-crisis levels even at the end of June, 2020.

2.4.1 A theoretical framework

There are two types of agents: a measure N of customers and a measure one of dealers, all of whom are price takers. Each customer seeks to trade one share of an asset. We do not distinguish between purchases and sales; this simplification allows us to study the determinants of transaction costs, though it is worth noting that our model is silent on the determinants of the asset's price. Since there are N customers with unit demand, the aggregate demand for transactions is exogenous and equal to N. However, while the total number of transactions is exogenous, the composition is not. Namely, we assume that customers demand vertically differentiated transaction services supplied by dealers at a convex cost: low-quality (l) transaction services, interpreted as agency trades, and high-quality (h) transaction services, interpreted as risky-principal trades.

Customers have quasi-linear utility for transaction services and for cash. Specifically, the problem of a customer is to choose how much low- and high-quality transaction services to demand from dealers at each time t in order to solve

$$\max_{x_{lt}, x_{ht}} u(x_{lt}, x_{ht}) + \theta_t x_{ht} - p_{lt} x_{lt} - p_{ht} x_{ht} \tag{2.5}$$

$$\text{sub. to} \quad x_{lt} + x_{ht} = 1. \tag{2.6}$$

We assume that $u(x_{lt}, x_{ht})$ is increasing, concave, twice continuously differentiable, and satisfies $u_h(x_{lt}, x_{ht}) - u_l(x_{lt}, x_{ht}) \geq 0$, where the h and l subscripts denote first partial derivatives with respect to x_{ht} and x_{lt}, respectively. This condition simply means that the customer values high-quality transaction services more than low-quality transaction services. The shock θ_t in the objective captures time variation in customer's utility for upgrading agency trades into risky-principal trades, or what we call their demand for immediacy.[24]

Assuming interior solutions, the first-order optimality conditions can be written

$$p_{jt} = u_j(x_{lt}, x_{ht}) - \lambda_t + \mathbf{1}_{\{j=h\}}\theta_t, \qquad j \in \{l, h\}, \tag{2.7}$$

for some multiplier λ_t on the constraint $x_{lt} + x_{ht} = 1$.

On the other side of the market, dealers choose their supply of transaction services, X_{lt} and X_{ht}, in order to maximize profits,

$$p_{lt}X_{lt} + p_{ht}X_{ht} - C(X_{lt}, X_{ht}),$$

where $C(X_{lt}, X_{ht})$ is some continuous, convex, and twice continuously differentiable cost function. This leads to the first-order optimality conditions

$$p_{jt} = C_j(X_{lt}, X_{ht}), \qquad j \in \{l, h\}. \tag{2.8}$$

[24] One could also add a time varying constant to the objective, capturing time variation in the utility for all transaction services, agency or risky principal. But since this constant would not appear in the first-order conditions, it would not change the demand analysis below.

Finally, the market clearing conditions for transaction services are simply

$$X_{jt} = N_t x_{jt}, \qquad j \in \{l, h\}. \tag{2.9}$$

An equilibrium is thus described by a sequence $\{x^\star_{lt}, x^\star_{ht}, X^\star_{lt}, X^\star_{ht}, p^\star_{lt}, p^\star_{ht}\}$ satisfying equations (2.6)–(2.9) at each time t.

2.4.2 Comparative statics

Combining the assumption of fixed-size demand, (2.6), with the customers' first-order optimality conditions in (2.7), we can express customers' *demand for immediacy* as a single equation in two unknowns, $(p_{ht} - p_{lt})$ and x_{ht}:

$$p_{ht} - p_{lt} = u_h(1 - x_{ht}, x_{ht}) - u_l(1 - x_{ht}, x_{ht}) + \theta_t. \tag{2.10}$$

This equation defines the inverse demand for immediacy: the relationship between the price premium for risky-principal trades and customers' marginal utility for upgrading from slow, agency trades to fast, risky-principal trades. As anticipated above, θ_t is a demand shock that generates a parallel shift of this inverse demand curve.

Exploiting the market clearing conditions in (2.9), in conjunction with (2.6) and (2.8), similar steps reveal an equation that captures the dealers' *willingness to supply immediacy*:

$$p_{ht} - p_{lt} = C_h\left(N_t(1 - x_{ht}), N_t x_{ht}\right) - C_l\left(N_t(1 - x_{ht}), N_t x_{ht}\right). \tag{2.11}$$

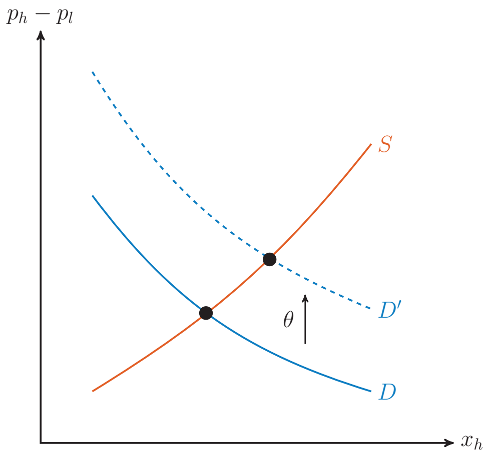
Hence, the equilibrium characterization reduces to a price premium, $p_{ht} - p_{lt}$, and a fraction of risky-principal trades, x_{ht}, that lies at the intersection of these demand and supply schedules.

This simple representation offers a parsimonious, transparent framework to analyze the effects of various types of shocks. First, shocks to consumers' relative preference for immediate, risky-principal trades, as captured by θ_t, shift the demand for immediacy but not the supply. As is evident from panel (a) of Figure 2.4.1, a positive innovation to θ_t induces an increase in the relative price of risky-principal trades, along with an increase in the equilibrium fraction of such trades.

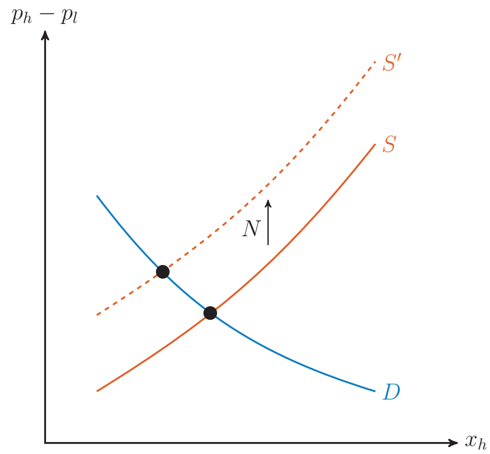
Alternatively, under natural conditions, a surge in customer-to-dealer trading volume, as captured by an increase in N_t, has no effect on each customer's demand for immediacy, but causes an upward shift in the supply curve, as providing risky-principal transaction services becomes more costly as the total volume of transaction services grows. Of course, any shock to the cost function $C(\cdot,\cdot)$—perhaps due to an increase in intermediaries' risk aversion or cost of funding—would engender a similar shift in the supply curve. As is evident from panel (b) of Figure 2.4.1, an upward shift in dealers' supply curve causes an increase in the price premium paid for risky-principal trades, but a decrease in the equilibrium fraction of such trades.

2.4.3 Estimating the model

In the data, we observed, simultaneously, an increase in the price premium $p_{ht}-p_{lt}$ (Figure 2.3.1) and a decrease in the fraction of risky-principal trade x_{ht} (Figure 2.3.2). According to the model, this is indicative of a supply shock. As noted above, a supply shock could have been generated by an increase in the total volume of transaction services which, when combined with binding balance sheet constraints,



(a) Shock to customers' preferences

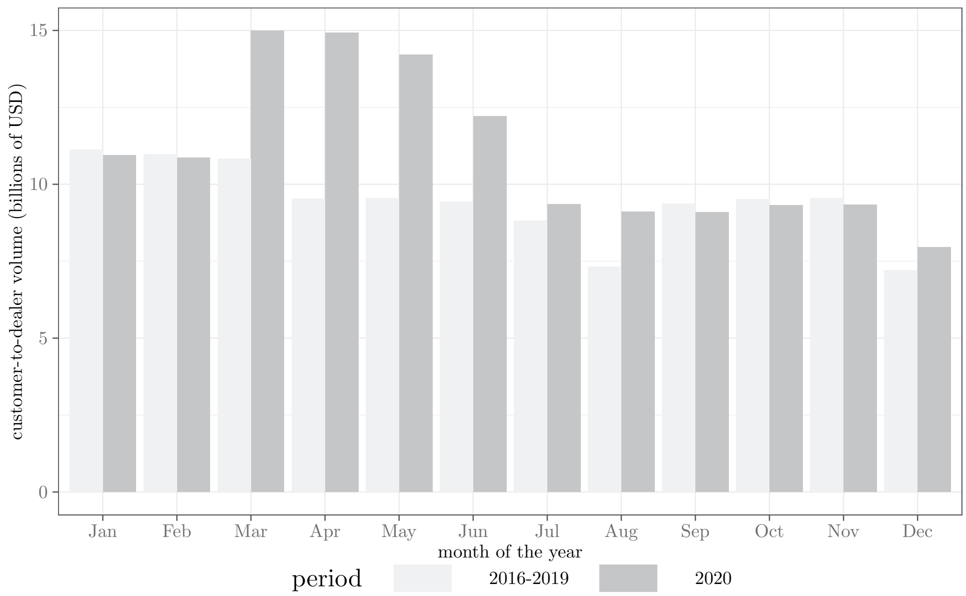


(b) Shock to dealers' cost

2.4.1. Demand and supply shocks.

would make it more costly for dealers to supply risky-principal trades. Indeed, Figure 2.4.2 illustrates that customer-to-dealer volume was about 50 percent higher during the crisis, relative to the (average in the) same months in 2016-2019. The figure also shows that the increase in volume was quite persistent, remaining above normal levels through the summer.

However, the data does not rule out shocks to the demand for immediacy (θ_t), as it is entirely possible that both the inverse supply and demand curves shifted in the same direction. To separately identify demand from supply shocks, we proceed in two steps. In this section, after imposing a specific functional form on customers' preferences, we estimate the parameter that determines the shape of the inverse demand curve by exploiting shifts to the supply curve that occurred during periods *outside of the crisis*, i.e., in "normal" times. Then, in Section 2.4.4, we use our estimated inverse demand curve to decompose movements in the price premium and the fraction of risky-principal trades *during the crisis* into movements that occur



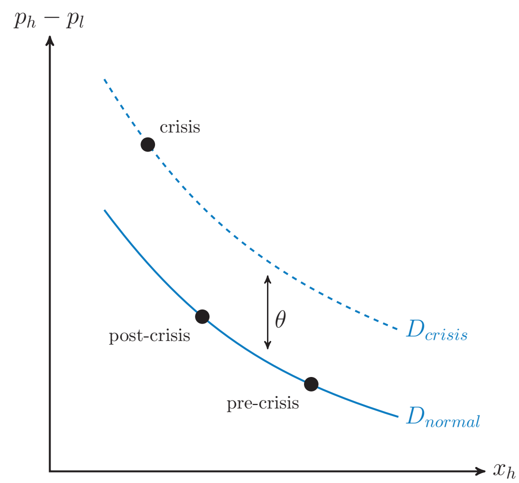
2.4.2. Customer-to-dealer volume. This figure plots the monthly volume of customer-to-dealer trades for 2020 and the 2016–2019 average in billions of USD.

along the demand curve—caused by shocks to the cost function, including innovations to N_t—and movements caused by innovations to θ_t, which shift the demand curve.

We plot this basic intuition in Figure 2.4.3. Note that this strategy is, by construction, largely independent of the shape of the supply curve and the nature of the shocks that shift it. However, by separately identifying shocks to the demand for immediacy from shocks to supply, we can study how the two responded differently to policy, and the quantitative implications for consumer surplus from immediacy throughout the pandemic, which we do in Sections 2.4.4 and 2.4.5, respectively.

2.4.3.1 Parametric specification

Since x_h and x_l represent the market shares of high- and low-quality transactions in a vertically differentiated market, respectively, a natural choice for the demand



2.4.3. Identifying relative preference shocks, θ.

curve is a logit specification.[25] In particular, we assume that, for each dollar of transaction service, the utility function of a consumer is given by

$$\theta_t x_{ht} - \sigma \left[x_{lt} \log(x_{lt}) + x_{ht} \log(x_{ht}) \right]. \tag{2.12}$$

It is well known (see, e.g., page 77 of Anderson et al., 1992) that this specification is equivalent to assuming that, for each dollar of transaction services, a consumer chooses between agency and risky-principal trades with net utilities $1 - p_{lt} + \varepsilon_{lt}$ and $1 + \theta_t - p_{ht} + \varepsilon_{ht}$, respectively, where ε_{lt} and ε_{ht} are independently and identically distributed (IID) over time and across consumers according to Gumbel distribution with location parameter zero and scale parameter σ.

Given this parametric assumption, the inverse demand for immediacy takes a

[25] Classic references for this specification include McFadden (1973), Anderson et al. (1992), and Berry (1994).

log-linear form:

$$p_{ht} - p_{lt} = -\sigma \log\left(x_{ht}/x_{lt}\right) + \theta_t. \tag{2.13}$$

As one would expect, a larger price premium, $p_{ht} - p_{lt}$, results in a lower demand for risky-principal trades, x_{ht}. In addition, one sees that the shape of the demand curve depends on just one semi-elasticity parameter, σ.

As is well known, a simple OLS regression of the price premium $p_{ht} - p_{lt}$ on log quantities, $\log(x_{ht}/x_{lt})$, would yield a biased estimate of σ, since relative quantities are, in general, correlated with the relative demand shock θ_t. Hence, we use an instrumental variable (IV) approach to estimate the parameter of interest, σ.

To do so, consider an arbitrary instrument Z_t for the log relative quantities $\log(x_{ht}/x_{lt})$. Using the inverse demand specification in (2.13), one easily sees that

$$\beta_{IV} = -\frac{\operatorname{Cov}\left(Z_t, p_{ht} - p_{lt}\right)}{\operatorname{Cov}\left(Z_t, \log(x_{ht}/x_{lt})\right)} = \sigma - \frac{\operatorname{Cov}\left(Z_t, \theta_t\right)}{\operatorname{Cov}\left(Z_t, \log(x_{ht}/x_{lt})\right)}. \tag{2.14}$$

Hence, as is well known, β_{IV} is an unbiased estimator of σ when the instrument Z_t is uncorrelated with the demand shock θ_t.

A binary IV approach

Consider observations about prices and relative quantities in two periods: a pre-crisis period, such as January 2020; and a post-crisis period, such as June 2020. Suppose the instrument Z_t takes the value zero in the pre-crisis period and 1 in the post-crisis period. Since

$$\operatorname{Cov}(Z_t, \theta_t) = \Pr(Z_t = 1)(1 - \Pr(Z_t = 1))\left(\mathbb{E}\left[\theta_t \mid Z_t = 1\right] - \mathbb{E}\left[\theta_t \mid Z_t = 0\right]\right),$$

it follows that an IV estimate based on Z_t is consistent if $\mathbb{E}\left[\theta_t \mid Z_t = 1\right] = \mathbb{E}\left[\theta_t \mid Z_t = 0\right]$. In other words, as long as the relative demand shock in the post-crisis period has returned to its pre-crisis average, then it is uncorrelated with the binary instrument, Z_t. Of course we also need the binary instrument to be relevant, i.e., correlated with the relative quantities. However, this is verified empirically because, as shown in Figure 2.3.2, relative quantities in the post-crisis period are lower than in the pre-crisis period. We interpret this observation as follows: as shown in Figure 2.4.2, trading volume in June remained elevated relative to pre-pandemic levels, which, in our model, shifts the marginal cost of providing transaction services, creating a supply shock. In equilibrium, the relative quantities demanded by consumers are reduced along a fixed demand curve.

The binary IV approach leads to the following candidate estimate

$$\hat{\sigma} = -\frac{\frac{1}{T_1}\sum_{t\,:\,Z_t=1}(p_{ht} - p_{lt}) - \frac{1}{T_0}\sum_{t\,:\,Z_t=0}(p_{ht} - p_{lt})}{\frac{1}{T_1}\sum_{t\,:\,Z_t=1}\log(x_{ht}/x_{lt}) - \frac{1}{T_0}\sum_{t\,:\,Z_t=0}\log(x_{ht}/x_{lt})},$$

where T_0 and T_1 are the lengths of pre- and post-crisis periods in days, respectively. For the estimation, we set the pre-crisis period to run between January 15, 2020 and February 14, 2020, and the post-crisis period to run between June 1, 2020 and June 30, 2020. We obtain an estimate of $\hat{\sigma} = 100.09$ with a standard deviation of 15.40.[26]

[26]Our assumption that θ_t returned to pre-crisis levels by June 2020 is supported by a variety of empirical evidence. For example, consistent with our own findings, Haddad et al. (2020) document that the severe dislocations in the corporate bond market disappeared soon after the Fed's interventions. Falato et al. (2020) and Ma et al. (2020) offer further evidence, establishing that bond mutual funds and ETFs experienced record *inflows* in April and May after massive outflows in March 2020. However, for robustness, we perform a second binary IV estimation that does not rely on the assumption that $\mathbb{E}_{\text{Jan}}\left[\theta_t\right] = \mathbb{E}_{\text{Jun}}\left[\theta_t\right]$, but rather that θ_t had settled down to *some* (arbitrary)

A high-frequency IV approach

An alternative approach to estimating σ is to use high frequency variation in trading conditions that affect prices and quantities but are unlikely to be attributable to aggregate shocks to customers' preferences for risky-principal trades. According to our theory, changes in N_t—which could equivalently represent the total volume of trade or the total number of trades (since trade size is fixed)—change the dealers' cost of supplying transaction services, but do not change individual consumers' relative demand for risky-principal trades.[27]

Hence, we consider two instruments to measure deviations in the quantity of customer-dealer trading: one based on trading volume and the other based on the number of trades in each day. Importantly, we exclude the crisis period March 1, 2020 to April 15, 2020, as shocks to relative demand during this period are likely significant and correlated with changes in both measures of N_t. We seasonally-adjust and detrend (by adding month dummies) $\log(N_t)$ for both measures, so that shocks represent the residual deviation. After constructing these series, the formal exclusion restriction for the IV estimate to be consistent is $\text{Cov}\left(\log(N_t), \theta_t\right) = 0$. Relative to the first binary IV, this approach has one advantage: it does not assume that the

level by May 2020. In particular, we assume that the average value of θ_t in the first half of May is equal to the average value in the second half of June, and repeat a similar estimation procedure to the one described above. We arrive at an estimate of $\hat{\sigma} = 80.06$, within 1.3 standard deviations of our initial binary IV approach. We provide more details about both binary IV estimations in Appendix 2.D.

[27]Intuitively, one could imagine that dealers cannot fully adjust the size of their balance sheets or their trading infrastructure in the short term (our time period is a day). Thus, in the short run, an increase in the number of market participants or trading volume could put upward pressure on dealers' cost of providing immediacy, changing the relative prices of risky-principal and agency trades without shifting relative demand. Importantly, note that we are assuming that innovations to N_t are uncorrelated with customers' *relative* demand for risky-principal trades, and not with customers' *aggregate* demand for transaction services.

2.4.1. Estimating the logit demand parameter σ using high-frequency IV approach. This table presents the IV estimates of the logit demand parameter σ. In column (1), the seasonally-adjusted log volume of trades is used as an instrument. In column (3), the seasonally-adjusted log number of trades is used as an instrument. Standard error are the maximum of robust and the usual standard errors. The pre-crisis runs from January 3, 2020 until February 29, 2020. The post-crisis data begins on April 15, 2020 and runs until July 31, 2020. We exclude holidays, weekends and half trading days. Our estimates and standard errors are transformed using the delta-method where appropriate. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

	Dependent variable:	
	$(p_h - p_l)$	
	IV (vol)	IV (num)
	(1)	(2)
$\log(x_h/x_l)$	70.17**	73.31**
	(33.63)	(31.02)
Post-crisis	8.25	7.75
	(5.75)	(5.14)
Constant	75.55***	78.15***
	(27.94)	(25.79)
Observations	113	113
Adjusted R^2	.41	.39

relative demand shock θ_t has returned to normal in June.

Table 2.4.1 presents the estimates of σ that emerge from our high frequency IV. The estimates range from 70 to 73 depending on the instrument, falling near the lower bound of the confidence interval of the binary-IV estimates.[28]

[28]As a robustness check, in Appendix 2.D, we combine both our instruments and estimate the overidentified system using two-stage least square (2SLS). We also provide first-stage results for each IV regression verifying that each one is a valid instrument. For our 2SLS where we use both number and volume as instruments, we provide test statistics for the Sargan-Hansen test of overidentifying restrictions.

2.4.4 Demand Shocks, Supply Shocks, and Policy Implications

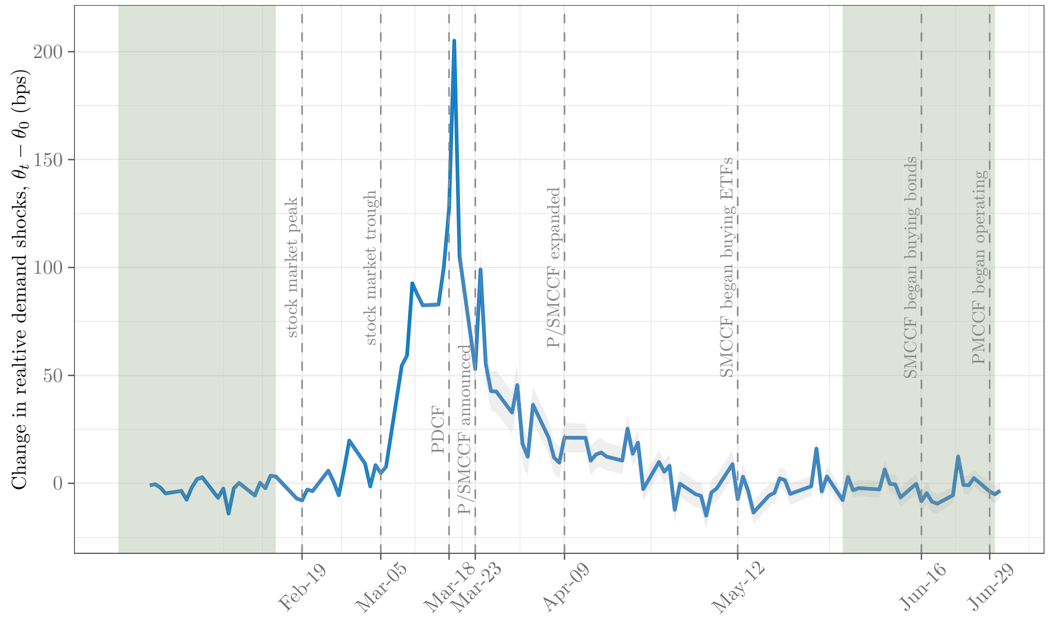
In this section, we discuss our model's (qualitative and quantitative) implications for the relative importance of shocks to customers' demand for immediacy and dealers' willingness to supply it over the course of the COVID-19 crisis and the interventions that followed.

To start, given an estimate of the semi-elasticity parameter σ, along with the time series for the price premium $(p_{ht} - p_{lt})$ and the ratio of agency trades to risky-principal trades (x_{lt}/x_{ht}), we can infer the sequence of shocks to customers' relative demand for risky-principal trades (θ_t). Using our estimate of σ from the binary IV described above, $\hat{\sigma} = 100.09$, Figure 2.4.4 plots the time series relative to a pre-crisis benchmark, $\theta_t - \theta_0$, where θ_0 is the inferred value of θ on January 2, 2020.[29]

Not surprisingly, the figure reveals that θ_t experiences a dramatic increase during the most tumultuous weeks of March. In fact, our estimates suggest that customers' willingness to pay for each inframarginal unit of risky-principal trade (rather than an agency trade) increased by approximately 200 bps at the height of the crisis, before receding quickly after the announcements of the Fed's interventions. To give a sense of the effect of this shock, note that we can decompose the change in the price premium if we assume that supply is perfectly inelastic:

$$p_{ht} - p_{lt} - (p_{h0} - p_{l0}) = \theta_t - \theta_0 + \sigma \left[\log\left(x_{lt}/x_{ht}\right) - \log\left(x_{l0}/x_{h0}\right)\right]. \tag{2.15}$$

[29]Our results are largely independent of which estimate we use, as is clear from the error bands in the plots. This is because the estimate of $\hat{\sigma}$ from the high-frequency IV estimation lies within two standard deviations of the estimate from the binary IV, and because log quantities vary less than the price premium.



2.4.4. The change in the estimated relative demand shocks for risky-principal trades. This figure plots the time-series of the change in the estimated relative demand shocks for risky-principal trades relative to a pre-crisis benchmark on January 2, 2020, $\theta_t - \theta_0$, implied from Equation (2.13). Plus/minus two standard deviation bands are shown in gray. The shaded areas correspond to the early and late periods used to estimate the semi-elasticity parameter σ.

In general, this decomposition depends on the relative elasticity of demand and supply. However, we believe that the case of a perfectly inelastic supply is a natural benchmark, for two reasons. First, it delivers an upper (lower) bound for the contribution of demand (supply) shocks to the price premium. Second, the assumption of an inelastic supply accords with reports of binding balance-sheet constraint at the height of the crisis, and is seemingly confirmed by our evidence on dealers' inventory accumulation from Figure 2.3.3.

According to this decomposition, a large portion of the spike in the cost of risky-principal trades (relative to agency trades) at the height of the crisis can be explained by relative demand shocks. For example, the average value of $p_{ht} - p_{lt}$ over the time period March 5-April 9 was 75 bps higher than the value on January 2. The portion of this increase in the price premium that can be explained by the demand shifter was 58 bps, or approximately 75%. Hence, our results are consistent with other studies that highlight the "dash for cash" as an important driver of the turmoil in the corporate bond market (such as Falato et al., 2020; Ma et al., 2020; Haddad et al., 2020).

However, an increase in the demand for immediacy alone would generate an *increase* in x_h, as in Figure 2.4.1, which is opposite of what we observe in the data. Therefore, a second key takeaway from our analysis is that the onset of the pandemic must have also induced a negative shock to dealers' willingness to use their balance sheet space to accommodate the surge in selling pressure. Indeed, within the context of Figure 2.4.3, our results suggest that the relative supply of risky-principal trades would have to experience a significant shift to the left in order to induce a drop in x_h. Thus, in addition to a surge in customers' demand for immediacy, our analysis reveals an equally important shock to dealers' costs of supplying immediacy; simply put, matching the data on prices *and quantities* requires an increase in the expected cost of dealers adding inventory to their balance sheets, as documented in the Treasury market by He et al. (2020).[30]

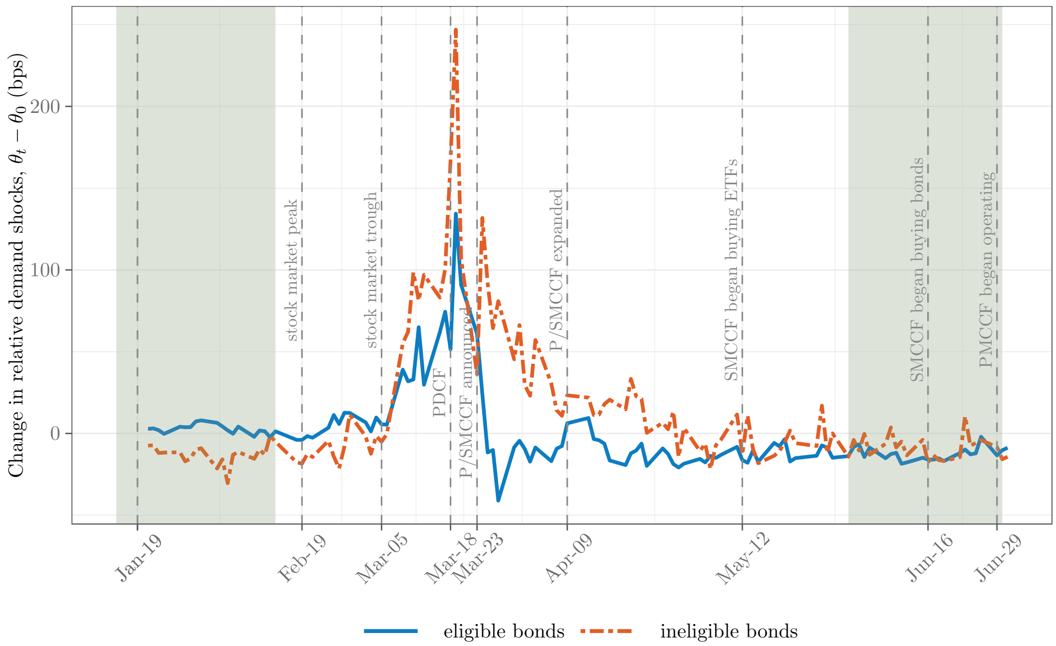
Finally, studying the behavior of the series $\{\theta_t, p_{ht} - p_{lt}, x_{ht}\}$ against the timeline of the Fed's interventions reveals important, new insights into the channels through which various policies impacted market liquidity. To start, the time path of relative

[30] Studying the market for mortgage-backed securities (MBS), Chen et al. (2020) also find that the combined liquidity constraints of customers and dealers are responsible for the severe price dislocations observed during the COVID-19 pandemic.

preference shocks suggests that the *announcement* of the Fed's interventions was enough to halt and reverse the "dash for cash" that began in the second week of March. To explore the relationship between customers' preferences and the Fed's interventions in greater depth, we can estimate the time path of θ_t separately for those bonds that were eligible vs. ineligible for purchase according to the March 23 announcement of the corporate credit facilities; Figure 2.4.5 plots the results.[31] The figure reveals that the Fed's announcement had a more immediate impact on the relative demand for risky-principal trades of eligible bonds, relative to ineligible bonds. Taken together with the results in Table 2.3.4, it appears that the expectation of price support from the Fed played an important role in halting the rush among customers to liquidate corporate debt immediately, and thus helped to reduce bid-ask spreads for risky-principal trades.

After the dash for cash subsided, and the demand curve shifted down, we observe that x_h does not decrease, but rather increases slightly. Hence, it must be that the announced interventions also reduced dealers' perceived cost of supplying risky-principal trades, shifting their supply curve to the right and inducing them to absorb inventory onto their balance sheet, as documented in Figure 2.3.3. However, the fact that the price premium and the fraction of agency trades remained elevated, even months after the initial shock appears to have passed, suggests that the supply curve did not return to its original location, i.e., that balance sheet costs remained higher than pre-crisis levels despite calmer markets and the Fed's interventions. These costs could derive from persistently high trading volume (as documented in Figure 2.4.2), from expectations of future price declines or volatility, or from losses incurred on

[31]We provide more details about estimated preference shocks for eligible and ineligible bonds in Appendix 2.E.



2.4.5. The change in the estimated relative demand shocks for risky-principal trades for eligible and ineligible bonds. This figure plots the time-series of the change in the estimated relative demand shocks for risky-principal trades relative to a pre-crisis benchmark on January 2, 2020, $\theta_t - \theta_0$, implied from Equation (2.13), for eligible and ineligible bonds. The shaded areas correspond to the early and late periods used to estimate the semi-elasticity parameter σ.

other parts of the dealers' balance sheets.

2.4.5 The Surplus from Immediacy

The analysis above highlights the important distinction between risky-principal and agency trades during the COVID-19 crisis: we have shown that customers' demand for immediacy increased drastically at the height of the crisis, while dealers' willingness to supply immediacy by accumulating inventory on their balance sheets simultaneously declined. In this section, we attempt to quantify the effects of these pandemic-induced shocks—and the interventions that followed—on the net utility that customers derived from immediacy.[32]

2.4.5.1 Theory

Given equilibrium prices and allocations, we define the consumer surplus from immediacy, per dollar of transaction, as

$$s_{ht} = \theta_t x^{\star}_{ht} + u(x^{\star}_{lt}, x^{\star}_{ht}) - (p^{\star}_{ht} - p^{\star}_{lt})\, x^{\star}_{ht}.$$

For each dollar of transaction, s_{ht} measures the customers' extra value of upgrading a fraction x_{ht} of agency trades into risky-principal trades, at a cost $(p_{ht} - p_{lt})\, x_{ht}$.[33]

[32]Note that we are intentionally not making any statements about optimal policy interventions or design. As we describe in greater detail below, our definition of consumer surplus from immediacy does not account for possible aggregate shocks to customers' demand for transaction services, nor does it capture any effects of prices and allocations in the corporate bond market on real investment decisions, possible linkages between corporate bond market liquidity and other funding markets, and so on. While certainly interesting, these extensions are beyond the scope of the current paper.

[33]We could also study the total surplus from all transaction, $s_{ht} - p_{lt}$. However, this measure could be biased, since our analysis of customer demand only identifies shocks to the demand for upgrading from risky-principal to agency trades, and not shocks to the overall demand for transaction services. There are at least two reasons to analyze the surplus from immediacy, as opposed to the total surplus. First, the bulk of the variation in transaction costs during the crisis, $p_{lt}x_{lt} + p_{ht}x_{ht}$, is accounted for by the cost of immediacy, $(p_{ht} - p_{lt})x_{ht}$. Second, much of the policy discussion surrounding the crisis was focused on the demand and supply for immediacy: customers' need to trade quickly and dealers' willingness to accommodate their demand with risky-principal trades that would have used their balance sheet space.

Namely, based on (2.6) and (2.10), one obtains:

$$\begin{aligned}
s_{ht} &= u(1 - x^\star_{ht}, x^\star_{ht}) - x^\star_{ht}\left[u_h(1 - x^\star_{ht}, x^\star_{ht}) - u_l(1 - x^\star_{ht}, x^\star_{ht})\right] \\
&= \int_0^{x^\star_{ht}} \left[u_h(1-y,y) - u_l(1-y,y)\right] dy - x^\star_{ht}\left[u_h(1 - x^\star_{ht}, x^\star_{ht}) - u_l(1 - x^\star_{ht}, x^\star_{ht})\right] \\
&= -\int_0^{x^\star_{ht}} \left[u_{hh}(1-y,y) - 2u_{lh}(1-y,y) + u_{ll}(1-y,y)\right] y\, dy, \qquad (2.16)
\end{aligned}$$

where the final equality follows from integration by parts. The term $-\left[u_{hh} - 2u_{lh} + u_{ll}\right]$ in the integral represents the slope of the inverse demand curve. Hence, the integral measures the area between the price premium and the inverse demand curve and so captures the surplus from upgrading from low-quality to high-quality transaction services.

Notice that the demand shock θ_t does not appear in the surplus from immediacy (2.16). Indeed, for any fixed $x^\star_{ht}$, the surplus from immediacy is the same regardless of the location of demand. This is because a parallel shift to the inverse demand curve increases the willingness to pay for all infra-marginal units by the same amount, θ_t. Put differently, when θ_t increases but $x^\star_{ht}$ stays the same, customers derive more utility from immediacy but pay more for it, with zero net effect on their surplus. Of course, in equilibrium, $x^\star_{ht}$ does not stay the same: hence the change in surplus from immediacy between time zero and time t is the consequence of customers substituting from risky-principal to agency trades:

$$s_{ht} - s_{h0} = -\int_{x^\star_{h0}}^{x^\star_{ht}} \left[u_{hh}(1-y,y) - 2u_{lh}(1-y,y) + u_{ll}(1-y,y)\right] y\, dy.$$

In other words, what ultimately makes consumers worse or better off is the net effect that the supply and demand shocks have on the quantity of risky-principal trades consumed in equilibrium.

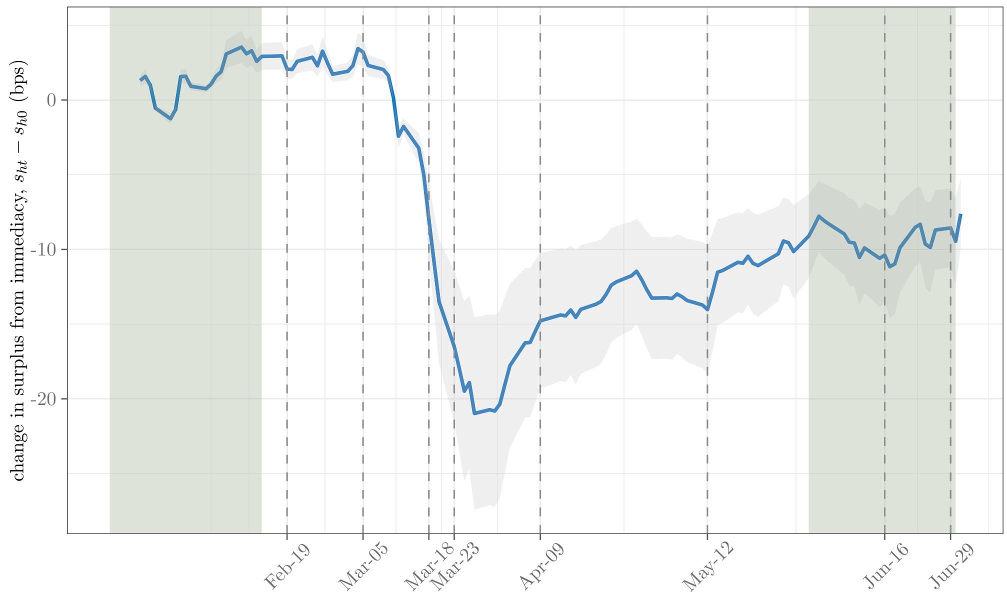
With the logit specification, $-(u_{hh} - 2u_{lh} + u_{ll}) = 1/[y(1-y)]$ and so we obtain a simple, closed-form expression for the change in the surplus from immediacy between time zero and time t:

$$s_{ht} - s_{h0} = -\sigma \log\left(\frac{1 - x_{ht}^{\star}}{1 - x_{h0}^{\star}}\right). \tag{2.17}$$

2.4.5.2 Estimate of surplus from immediacy

Using Equation (2.17) together with our estimate of σ, Figure 2.4.6 plots the change in consumer surplus from immediacy, per unit of transaction, over time. The figure reveals, not surprisingly, that there was a sharp, significant decline at the height of the market turmoil in mid-March, 2020. However, the figure also reveals that this decline was persistent: consumer surplus from immediacy per unit of transaction remained approximately 10 bps below pre-crisis levels even at the end June.

Comparing the dynamics of the surplus from immediacy to the expenditures on immediacy, $(p_{ht} - p_{lt})x_{ht}$, reveals two important differences. First, one can easily confirm that the decline in the surplus from immediacy is much smaller than the increase in expenditures. Second, the recovery takes longer to materialize and is less dramatic. As explained above, these differences arise because the surplus from immediacy accounts for two additional effects that matter for evaluating consumer well-being. First, customers' preference for immediacy change: this implies that the surplus loss induced by the increase in immediacy expenditure is partly offset by the additional value derived from immediacy. Second, the composition of transaction services changes as well: consumers substitute towards agency trades, so that the average quality of transaction services they enjoy falls. We believe these observations



2.4.6. The change in consumer surplus from immediacy in the logit demand specification. This figure shows the change in surplus for immediacy relative to a pre-crisis benchmark on January 2, 2020, $s_{ht} - s_{h0}$, according to Equation (2.17). Plus/minus two standard deviation bands are shown in gray. The shaded areas correspond to the early and late periods used to estimate the semi-elasticity parameter σ.

serve as an important reminder that changes in consumers' well-being is often not well-approximated by changes in prices.

2.5 Conclusion

It often takes a bad shock to discover whether or not a market is liquid, and to expose any sources of illiquidity. Unfortunately, many shocks to financial markets originate *inside* financial intermediaries, and hence the aggregate shock *is* a liquidity shock. In this sense, the COVID-19 pandemic—a truly exogenous, large shock that did not originate in the banking sector—offers a unique opportunity to study market conditions, the shocks that precipitate episodes of illiquidity, and the implications for transaction costs, policy, and consumer surplus.

In this paper, we study trading conditions in the US corporate bond market from many angles as the COVID-19 pandemic unraveled. However, a key insight is that distinguishing between risky-principal and agency trades offers not only a more complete assessment of market conditions, but also a unique window into the sources of illiquidity and the efficacy of policy interventions in this large, important market. In particular, we find that the initial panic was caused by shocks to both customers' demand for immediacy *and* dealers' willingness to supply it. The former shock receded quickly, and almost fully, after the mere announcement of the Fed's intention to enter the market and purchase bonds. The latter shock, however, lingered months after markets appeared to calm, indicating that elevated trading volume in conjunction with balance sheet constraints remain a risk in times of crisis.

While this is an important first step, much work remains to be done. Perhaps most importantly, further examination of dealers' balance sheets and changes (or heterogeneity) in regulatory requirements could allow us to pinpoint the precise source of dealers' unwillingness to "lean against the wind" during times of crisis. Identifying and understanding these constraints would allow us to design better policies to balance the crucial trade-off between risk-taking and liquidity provision often at the heart of liquidity provision in financial markets. We leave this work for the future.

APPENDIX

2.A Data and Definitions

2.A.1 Data description

We use data from the Trade Reporting Compliance Engine (TRACE), made available by the Financial Industry Regulation Authority (FINRA). The raw TRACE data provides detailed information on all secondary market transactions self-reported by FINRA member dealers. These include bond's CUSIP, trade execution time and date, transaction price ($100 = par), the volume traded (in dollars of par), a buy/sell indicator, and flags for dealer-to-customer and inter-dealer trades. To construct our sample, we combine two versions of TRACE: the standard version (2020Q1), and the End-Of-Day version (2020Q2).

We first filter the report data following the procedure laid out in Dick-Nielsen (2014). We merge the resulting data set with the TRACE master file, which contains bond grade information, and with the Mergent Fixed Income Securities Database (FISD) to obtain bond fundamental characteristics. Following the bulk of the academic literature, we exclude bonds with optional characteristics, such as variable coupon, convertible, exchangeable, and puttable, as well as, asset-backed securities, and private placed instruments. Table 2.A.1 provides summary statistics for our sample.

In our empirical specifications, we exclude newly-issued securities (with age less than 90 days), as on-the-run bonds tend to trade differently than off-the-run securities. Since our sample only contains about 130 days, the age and time-to-maturity

2.A.1. Summary statistics. This table provides mean, standard deviation, median, 5th and 95th percentiles of the average daily number of trades and volume by counterparty type, proportion of agency trades, proportion of trades on IG bonds, proportion of trades on the bonds eligible for SMCCF, and daily average trading cost for risky-principal (CH) and agency trades (MIRC) for eligible and ineligible bonds respectively. "num" refers to number of trades, and the "vol" refers to volume of trades in par value. A bond is considered eligible for the SMCCF if it has an investment-grade rating and time-to-maturity of 5 years or less on the March 23 2020. Source: TRACE and FISD. The sample starts on January 3 and ends on June 30, 2020.

	Mean	SD	Q05	Q50	Q95
Daily num. all trades	51,908	8,871	39,642	52,237	67,005
Daily num. interdealer	20,972	3,746	16,073	20,963	27,282
Daily num. customer	30,936	5,278	23,622	31,137	38,949
Daily num. customer-bought	16,114	2,859	12,187	16,239	20,630
Daily num. customer-sold	14,822	2,952	11,202	14,497	19,752
Daily vol. all trades ($b)	10.38	1.90	7.38	10.53	13.25
Daily vol. interdealer ($b)	3.13	0.61	2.19	3.17	4.04
Daily vol. customer ($b)	7.25	1.39	5.18	7.19	9.48
Daily vol. customer-bought ($b)	3.75	0.66	2.83	3.75	4.79
Daily vol. customer-sold ($b)	3.50	0.80	2.35	3.45	4.94
Prop. agency (num)	0.54	0.03	0.50	0.53	0.59
Prop. agency (vol)	0.33	0.04	0.29	0.33	0.38
Prop. IG (num)	0.74	0.03	0.70	0.74	0.78
Prop. IG (vol)	0.82	0.03	0.79	0.82	0.86
Prop. eligible (num)	0.33	0.04	0.30	0.33	0.40
Prop. eligible (vol)	0.27	0.04	0.23	0.27	0.34
Daily avg. CH (bps)	51.30	35.21	22.38	40.66	123.29
Daily avg. CH of Eligible bonds (bps)	28.11	26.34	10.71	20.85	65.86
Daily avg. CH of ineligible bonds (bps)	61.32	41.78	27.50	47.67	147.32
Daily avg. MIRC (bps)	10.57	3.38	7.31	9.62	17.43
Daily avg. MIRC of Eligible bonds (bps)	5.35	2.78	2.91	4.43	11.45
Daily avg. MIRC of ineligible bonds (bps)	12.33	4.12	8.19	11.18	20.51

of a particular bond will vary little over time. Thus, we do not include the standard cross-sectional controls related to the bond's age or time-to-maturity. Furthermore, since we exclude newly-issued bonds, over time, the age (maturity) of any bond will increase (decrease) by one day each day. Thus, the average age (maturity) of our bonds will increase (decrease) monotonically over time, meaning these controls will also correlate with the time trends we are documenting.

We also distinguish between bonds that are eligible for the SMCCF and ineligible bonds. In Appendix 2.B, we present a detailed description of eligibility criteria for the SMCCF. We define a bond as eligible if it has investment-grade rating and time-to-maturity of five years or less on March 23, 2020, when the SMCCF was first announced. The eligibility criteria also state that the firm must be a US-domiciled corporation. Specifically, the Fed restricts its purchases to bonds where

> The issuer is a business that is created or organized in the United States or under the laws of the United States with significant operations in and a majority of its employees based in the United States.

This criterion leaves the Fed with a considerable degree of discretion. For instance, if a foreign-domiciled corporation uses a US subsidiary to issue dollar-dominated debt, our firm-level data identify the firm as non-US. We would then classify its bonds as foreign, making them ineligible for the SMCCF. However, under the Fed's definition of a US issuer, the bonds may be eligible for purchase. Using the Fed's SMCCF transaction-level disclosures, we find that in many cases, the holding firm of the security is a non-US entity.[34] One such example is British American Tobacco (BAT), a firm listed on the London Stock Exchange and domiciled in the UK. Our firm-level data correctly identifies this firm as foreign; however, its bonds were purchased by the Fed.[35] These bonds were issued by a US wholly-owned subsidiary of BAT, BAT Capital Corporation. Since this subsidiary is guaranteed and wholly-owned

[34]SMCCF transaction-specific disclosures are provided by the Federal Reserve, available at https://www.federalreserve.gov/monetarypolicy/smccf.htm.

[35]On July 10, 2020, the Fed reported that BAT's bonds were purchased as part of the SMCCF (CUSIP 05526DAZ8).

by BAT, it is very challenging to correctly classify these bonds as US-domiciled. We, therefore, do not use US vs. non-US as an SMCCF eligibility criterion in our regressions discussed below and focus only on US firms.

Moreover, we do not have access to the latest credit rating data for all bonds in our sample. For the sub-sample of bonds where the credit rating is available, we include a credit rating fixed effect to control for potentially time-invariant nature of bond credit ratings.

2.A.2 Dates highlighted in the figures

We choose the following dates to highlight in the figures with vertical, dashed lines:

January 19: beginning of the series, chosen to start the sample period one month before the stock market peak.

February 19 stock market peak.

March 5: beginning of extended fall in equity prices and rise in corporate credit spreads.

March 18: first day of trading after announcement of Primary Dealer Credit Facility (announced evening of March 17).

March 23: announcement of Primary and Secondary Market Corporate Credit Facilities.

April 9: expansion of PMCCF and SMCCF (in both size and scope).

May 12: the SMCCF began purchasing eligible ETFs.

June 16: the SMCCF began purchasing individual corporate bonds.

June 29: the PMCCF began operating.

2.A.3 Identifying agency trades

We define agency trades as two trades in a given bond with the same trade size that take place within 15 minutes of each other. For each bond, we divide its trading sample into three groups: customer-sell-to-dealer (C2D), dealer-sell-to-customer (D2C), and interdealer (D2D) trades. Our identification of agency trades includes the following steps:

1. We match each trade X in group C2D with a trade Y in group D2C that has the same trade size and happens within 15 minutes of X. If there are several trades in D2C satisfying these conditions, we choose the trade that takes place closest in time to X. The identified pair of agency trades is then (X, Y). After this step, we denote the collection of unmatched trades in C2D as u-C2D and that in D2C as u-D2C.

2. We match each trade in u-C2D with a trade in group D2D by the same algorithm. We then obtain a collection of unmatched trades in D2D, denoted by u-D2D.

3. We match each trade in u-D2D with one in u-D2C following the same algorithm.

4. We repeat steps 1–3 using all remaining unmatched trades in the three groups while relaxing the matching criteria. In each agency trade pair, we require the second trade to happen within 15 minutes of the first trade, but it can have a

smaller trade size than the first one. By doing so, we consider the situation in which dealers split the trade volumes when they behave as matchmakers.

5. Finally, within all the remaining unmatched trades after steps 1–4, we identify trades with field `remuneration == "C"` in TRACE (commission is included in the price) as agency trades, because, by FINRA's definition, broker-dealers receive commissions only when they intermediate agency trades.

2.B Corporate Credit Facilities

In March 23, 2020, the Federal Reserve Bank of New York established the Primary Market Corporate Credit Facility (PMCCF) and the secondary Market Corporate Credit Facility (SMCCF). The purpose of the PMCCF was to sustain funding for corporate debt while the SMCCF was meant to support liquidity in the corporate bond market. These corporate credit facilities were funded by a 75 billion dollars investment, to be leveraged up to 750 billion. The SMCCF started its purchases of ETFs on May 12 and of corporate bonds on June 16. The PMCCF started its operations on June 29. On December 31, 2020, the corporate credit facilities stopped their purchases.[36]

2.B.1 Bond eligibility criteria for the SMCCF

The Federal Reserve established eligibility criteria for the purchases of corporate bonds. We provide excerpts from the Fed's own communications that detail these

[36]For more details, see the Frequently Asked Questions for PMCCF and SMCCF from the New York Fed, available at https://www.newyorkfed.org/markets/primary-and-secondary-market-faq/corporate-credit-facility-faq.

conditions.[37]

Eligible individual corporate bonds: The Facility may purchase individual corporate bonds that, at the time of purchase by the Facility: (i) were issued by an eligible issuer; (ii) have a remaining maturity of 5 years or less; and (iii) were sold to the Facility by an eligible seller.

Eligible issuers for individual corporate bonds: To qualify as an eligible issuer of an eligible individual corporate bond, the issuer must satisfy the following conditions:

1. The issuer is a business that is created or organized in the United States or under the laws of the United States with significant operations in and a majority of its employees based in the United States.
2. The issuer was rated at least BBB−/Baa3 as of March 22, 2020, by a major nationally recognized statistical rating organization ("NRSRO"). If rated by multiple major NRSROs, the issuer must be rated at least BBB−/Baa3 by two or more NRSROs as of March 22, 2020.
   (a) An issuer that was rated at least BBB−/Baa3 as of March 22, 2020, but was subsequently downgraded, must be rated at least BB−/Ba3 as of the date on which the Facility makes a purchase. If rated by multiple major NRSROs, such an issuer must be rated at least BB−/Ba3 by two or more NRSROs at the time the Facility makes a purchase.

[37]Source: Secondary Market Corporate Credit Facility Term Sheet available from https://www.federalreserve.gov/newsevents/pressreleases/files/monetary20200728a1.pdf, last updated on July 28, 2020.

(b) In every case, issuer ratings are subject to review by the Federal Reserve.

3. The issuer is not an insured depository institution, depository institution holding company, or subsidiary of a depository institution holding company, as such terms are defined in the Dodd-Frank Act.
4. The issuer has not received specific support pursuant to the CARES Act or any subsequent federal legislation.
5. The issuer must satisfy the conflicts of interest requirements of section 4019 of the CARES Act.

2.C Additional Empirical Results

2.C.1 The fraction of agency trades

In Table 2.C.1, we repeat the OLS regression in column (1) of Table 2.3.2 but focusing only on bonds issued by US firms. In columns (2) and (3) we repeat the regression in column (1) restricting the sample to eligible and ineligible bonds, respectively. As before, a bond is considered eligible if it has an IG credit rating and remaining time-to-maturity of five years or less.

Results in column (1), for US bonds, are very similar to what shown in column (1) of Table 2.C.1 for all bonds. From columns (2) and (3), we observe that the shift towards agency trades was much more pronounced among bonds that were eligible for the Fed's purchasing program. The probability of an agency trade for a given eligible bond, on average, rose by approximately seven percentage points relative to the pre-crisis period. After the Fed interventions on March 23, this probability

2.C.1. Robustness: Probability of an agency trade for US bonds (OLS only). This table presents regression results for the following specification from: $\text{Agency}_{ijt} = \alpha_i + \alpha_s + \beta_1 \times \text{Crisis}_t + \beta_2 \times \text{Intervention}_t + \varepsilon_{ijt}$. The dependent variable, Agency_{ijt}, is an indicator variable that takes the value 1 if trade j for bond i on day t is an agency trade and 0 otherwise. Only US firms are included in the regression. Crisis_t and Intervention_t are dummies which take the value of 1 if day t falls into Crisis and Intervention sub-periods defined above. There are three trade size categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million. A bond is considered eligible if it has an investment-grade rating and time-to-maturity of 5 years or less on the March 23 2020. The sample starts on January 3 and ends on June 30, 2020. Clustered standard errors at the day and bond levels are shown in parentheses. *$p<.1$; **$p<.05$; ***$p<.01$.

	Dependent variable:		
	Probability of agency trade		
	All	Eligible	Ineligible
	(1)	(2)	(3)
Crisis	0.043***	0.068***	0.025***
	(0.010)	(0.014)	(0.008)
Intervention	0.027***	0.044***	0.016***
	(0.003)	(0.004)	(0.004)
Bond FE	Yes	Yes	Yes
Trade size category FE	Yes	Yes	Yes
Observations	5,770,765	2,337,519	3,433,246
Adjusted R^2	.104	.071	.123

decreased from the crisis period (by 200 bps) to five percentage points higher than the pre-crisis period. For ineligible bonds, in contrast, the probability of an agency trade rose by only 1.9 percentage points relative to the pre-crisis period and remained relatively unchanged after the Fed intervention.

2.C.2 Impact of Fed announcements

In this subsection, we present several robustness checks for the difference-in-differences (DID) results in Section 2.3.5.

2.C.2.1 Bonds close to the eligibility threshold for rating and maturity

First, in Table 2.C.2, we repeat the regressions in Table 2.3.4 but focusing only on bonds just above and below the SMCCF eligibility threshold for time-to-maturity (TTM): bonds with four to six years left to maturity.

Next, in Table 2.C.3, we repeat the regressions in Table 2.C.2 but adding the extra restriction that the bonds should be close to the IG-HY threshold. In particular, we only include bonds that in addition to having TTM of four, five and six years, are also rated at the bottom tier of investment-grade (BBB+/Baa1, BBB/Baa2, and BBB−/Baa3) or the top tier of high-yield (BB+/Ba1, BB/Ba2, and BB−/Ba3).

2.C.2.2 Trade costs for different trade size bins

Here we run the regressions in (2.3.4) but with the trades of a particular size category in a different regression. In Tables 2.C.4-2.C.6, we show that small and large trades are responsible for the entire liquidity improvement documented in Table 2.3.4: small trades (with par volume of \$100,000 or less) become much more liquid after the Fed's CCF announcements followed by large trades (with volume larger than \$1 million). Liquidity of odd-lot trades (with volume between \$100,000 and \$1 million) seem to be unaffected by the Fed's intervention. Curiously we fail to find an affect for Odd-lot trades. There is some empirical evidence, e.g., Feldhütter (2012), suggesting

2.C.2. DID robustness: only include bonds with 4 to 6 years left to maturity. This table presents regression results for the following DID specification from Equation (2.4): $y_{ijt} = \alpha_s + \alpha_k + \beta_1 \times \text{SMCCF}_t \times \text{Eligible}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{Eligible}_t + \gamma \times X_{i,t} + \varepsilon_{ijt}$. The dependent variables are measures of transactions costs for risky-principal and agency trades. SMCCF_t is a dummy that takes the value of 1 if day t falls between March 23 and April 9, and 0 otherwise. Eligible_t takes the value of 1 if the bond has an investment-grade rating and time-to-maturity of 5 years or less on the March 23 2020. X_{it} controls for log(Amt outstanding), log(Age), and log(Time-to-maturity): logs of bond's amount outstanding, years since bond issuance, and years to maturity, respectively. There are three trade size categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million. The sample begins on March 6 and ends on April 9, 2020. Only US firms, bonds with 4, 5, or 6 years left to maturity on the intervention date are included. Bonds that change credit grade are excluded. Clustered standard errors at the day and bond levels are shown in parentheses. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

	Dependent variable:							
	Risky-principal				Agency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SMCCF × Eligible	-93.26^{**}	-76.08^{***}	-61.24^{***}	-61.42^{***}	-7.45^{*}	-9.13^{**}	-3.38	-3.79
	(39.33)	(27.87)	(17.83)	(17.89)	(4.36)	(4.22)	(4.14)	(4.09)
SMCCF	13.88	-0.46	-9.05	-9.14	2.82	5.15^{**}	1.46	1.75
	(35.04)	(25.14)	(16.58)	(16.64)	(2.37)	(2.47)	(2.14)	(2.14)
Eligible	54.22	12.27			-0.77	11.87^{**}		
	(50.71)	(33.55)			(4.42)	(5.46)		
log(Amount outstanding)	-3.86	-23.05^{**}			-4.06^{***}	-1.64		
	(16.73)	(10.71)			(1.03)	(1.07)		
log(Time-to-maturity)	-66.38	-102.01			8.79	30.42^{*}		
	(134.08)	(90.03)			(17.08)	(17.69)		
log(Age)	28.46^{*}	31.43^{**}			0.99	2.55^{*}		
	(15.69)	(12.45)			(1.95)	(1.45)		
Trade size category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	No	No	Yes	Yes	No	No
Bond FE	No	No	Yes	Yes	No	No	Yes	Yes
Credit rating FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	30,743	30,430	30,744	30,430	9,182	9,004	9,183	9,004
Adjusted R^2	.03	.07	.20	.20	.11	.14	.29	.30

that trades with different sizes are affected differently by market turmoil.

2.C.3. DID robustness: only include bonds with 4 to 6 years left to maturity and rating close to the IG/HY threshold. This table presents regression results for the following DID specification from Equation (2.4): $y_{ijt} = \alpha_s + \alpha_k + \beta_1 \times \text{SMCCF}_t \times \text{Eligible}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{Eligible}_t + \gamma \times X_{i,t} + \varepsilon_{ijt}$. The dependent variables are measures of transactions costs for risky-principal and agency trades. SMCCF_t is a dummy that takes the value of 1 if day t falls between March 23 and April 9, and 0 otherwise. Eligible_t takes the value of 1 if the bond has an investment-grade rating and time-to-maturity of 5 years or less on the March 23 2020. X_{it} controls for log(Amt outstanding), log(Age), and log(Time-to-maturity): logs of bond's amount outstanding, years since bond issuance, and years to maturity, respectively. There are three trade size categories: less than \$100,000, between \$100,000 and \$1 million, and larger than \$1 million. The sample begins on March 6 and ends on April 9, 2020. Only US firms, bonds with 4, 5, or 6 years left to maturity that are rated at the bottom tier of IG (BBB+/Baa1, BBB/Baa2, and BBB−/Baa3) or the top tier of HY (BB+/Ba1, BB/Ba2, and BB−/Ba3) are included. Bonds that change credit grade are excluded. Clustered standard errors at the day and bond levels are shown in parentheses. $^*p<.1$; $^{**}p<.05$; $^{***}p<.01$.

	Dependent variable:							
	Risky-principal				Agency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SMCCF × Eligible	-94.92^{**}	-86.46^{**}	-73.52^{**}	-73.52^{**}	−2.82	−5.38	−0.16	−0.16
	(45.73)	(43.93)	(30.56)	(30.56)	(3.58)	(3.80)	(3.90)	(3.90)
SMCCF	46.47	37.41	16.30	16.30	2.54	5.14^{**}	2.41	2.41
	(29.20)	(24.24)	(17.65)	(17.65)	(1.89)	(2.04)	(2.34)	(2.34)
Eligible	63.68	46.64			−1.18	7.74		
	(46.38)	(52.37)			(4.19)	(7.49)		
log(Amount outstanding)	−9.03	−18.19			-3.94^{***}	-2.58^{*}		
	(21.30)	(16.18)			(1.47)	(1.45)		
log(Time-to-maturity)	−160.39	−129.63			40.68	45.95		
	(150.14)	(130.02)			(30.85)	(30.34)		
log(Age)	36.44	30.68^{*}			−3.29	−1.56		
	(22.44)	(18.18)			(3.10)	(2.31)		
Trade size category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	No	No	Yes	Yes	No	No
Bond FE	No	No	Yes	Yes	No	No	Yes	Yes
Credit rating FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	14,124	14,124	14,124	14,124	4,595	4,595	4,595	4,595
Adjusted R^2	.04	.05	.16	.16	.12	.13	.28	.28

2.C.4. DID robustness: only include trades with par volume < $100,000, i.e., micro trades. This table presents regression results for US firms for the following DID specification from Equation (2.4): $y_{ijt} = \alpha_s + \alpha_k + \beta_1 \times \text{SMCCF}_t \times \text{Eligible}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{Eligible}_t + \gamma \times X_{i,t} + \varepsilon_{ijt}$. The dependent variables are measures of transactions costs for risky-principal and agency trades. SMCCF_t is a dummy that takes the value of 1 if day t falls between March 23 and April 9, and 0 otherwise. Eligible_t takes the value of 1 if the bond has an investment-grade rating and time-to-maturity of 5 years or less on the March 23 2020. X_{it} controls for log(Amt outstanding), log(Age), and log(Time-to-maturity): logs of bond's amount outstanding, years since bond issuance, and years to maturity, respectively. The sample begins on March 6 and ends on April 9, 2020. Only trades that are less than $100,000 in par volume, i.e., micro trades, are included. Bonds that change credit grade are excluded. Clustered standard errors at the day and bond levels are shown in parentheses. *$p < .1$; **$p < .05$; ***$p < .01$.

	Dependent variable:							
	Risky-principal				Agency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SMCCF × Eligible	-77.67^{***}	-54.42^{***}	-58.51^{***}	-45.93^{***}	-16.21^{***}	-19.37^{***}	-14.86^{***}	-15.66^{***}
	(19.69)	(18.51)	(12.66)	(12.56)	(4.45)	(4.53)	(5.24)	(5.27)
SMCCF	3.05	-28.30	-18.24	-30.64	9.45^{***}	11.48^{***}	6.63^{***}	7.26^{***}
	(24.08)	(23.12)	(20.75)	(21.29)	(3.06)	(2.97)	(2.28)	(2.36)
Eligible	0.94	-22.17			6.48	15.67^{***}		
	(23.80)	(18.37)			(4.66)	(4.64)		
log(Amt outstanding)	-38.79^{***}	-32.77^{**}			-4.24^{***}	-2.03^{**}		
	(11.74)	(13.90)			(0.79)	(0.79)		
log(Time-to-maturity)	11.54^{*}	7.93			4.26^{***}	5.98^{***}		
	(6.90)	(7.99)			(1.08)	(1.58)		
log(Age)	39.02^{***}	37.74^{***}			7.44^{***}	8.23^{***}		
	(13.24)	(11.15)			(1.85)	(1.84)		
Industry FE	Yes	Yes	No	No	Yes	Yes	No	No
Bond FE	No	No	Yes	Yes	No	No	Yes	Yes
Credit rating FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	92,300	82,694	92,301	82,694	28,556	27,182	28,556	27,182
Adjusted R^2	.05	.08	.35	.37	.05	.08	.26	.27

2.C.5. DID robustness: only include trades with \$100,000 $\leq$ par volume < \$1 million, i.e., odd-lot trades. This table presents regression results for US firms for the following DID specification from Equation (2.4): $y_{ijt} = \alpha_s + \alpha_k + \beta_1 \times \text{SMCCF}_t \times \text{Eligible}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{Eligible}_t + \gamma \times X_{i,t} + \varepsilon_{ijt}$. The dependent variables are measures of transactions costs for risky-principal and agency trades. SMCCF_t is a dummy that takes the value of 1 if day t falls between March 23 and April 9, and 0 otherwise. Eligible_t takes the value of 1 if the bond has an investment-grade rating and time-to-maturity of 5 years or less on the March 23 2020. X_{it} controls for log(Amt outstanding), log(Age), and log(Time-to-maturity): logs of bond's amount outstanding, years since bond issuance, and years to maturity, respectively. The sample begins on March 6 and ends on April 9, 2020. Only trades that are greater than \$100,000 and less than \$1 million in par volume, i.e., odd-lot trades, are included. Bonds that change credit grade are excluded. Clustered standard errors at the day and bond levels are shown in parentheses. $^{*}p$ <.1; $^{**}p$ <.05; $^{***}p$ <.01.

	Dependent variable:							
	Risky-principal				Agency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SMCCF × Eligible	−11.67	−8.28	−6.37	−5.44	−1.21	−1.66	−1.84	−1.01
	(12.57)	(13.87)	(13.85)	(14.15)	(2.50)	(2.35)	(2.56)	(2.34)
SMCCF	-27.03^{*}	-31.91^{*}	-36.84^{**}	-37.76^{**}	2.42	2.82^{*}	2.58	1.74
	(15.14)	(16.45)	(15.07)	(15.25)	(1.78)	(1.51)	(2.00)	(1.82)
Eligible	−0.33	−4.16			0.23	2.10		
	(12.17)	(13.48)			(1.99)	(1.81)		
log(Amt outstanding)	-18.54^{***}	-25.29^{***}			-2.97^{***}	-2.40^{**}		
	(4.34)	(3.84)			(0.99)	(1.02)		
log(Time-to-maturity)	26.48^{***}	33.45^{***}			3.86^{***}	3.00^{***}		
	(2.85)	(3.26)			(0.60)	(0.54)		
log(Age)	15.49^{***}	17.94^{***}			3.02^{***}	2.43^{***}		
	(3.10)	(3.08)			(0.58)	(0.76)		
Industry FE	Yes	Yes	No	No	Yes	Yes	No	No
Bond FE	No	No	Yes	Yes	No	No	Yes	Yes
Credit rating FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	36,406	34,457	36,407	34,457	10,089	9,775	10,089	9,775
Adjusted R^2	.03	.03	.07	.07	.04	.03	.20	.22

2.D Estimation Details

2.D.1 Binary IV

In this section, we provide more details about the binary IV estimation method for σ discussed in Section 2.4.3.1. Let's consider a linear supply and demand system

2.C.6. DID robustness: only include trades with par volume $\geq$ \$1 million, i.e., large trades. This table presents regression results for US firms for the following DID specification from Equation (2.4): $y_{ijt} = \alpha_s + \alpha_k + \beta_1 \times \text{SMCCF}_t \times \text{Eligible}_t + \beta_2 \times \text{SMCCF}_t + \beta_3 \times \text{Eligible}_t + \gamma \times X_{i,t} + \varepsilon_{ijt}$. The dependent variables are measures of transactions costs for risky-principal and agency trades. SMCCF_t is a dummy that takes the value of 1 if day t falls between March 23 and April 9, and 0 otherwise. Eligible_t takes the value of 1 if the bond has an investment-grade rating and time-to-maturity of 5 years or less on the March 23 2020. X_{it} controls for log(Amt outstanding), log(Age), and log(Time-to-maturity): logs of bond's amount outstanding, years since bond issuance, and years to maturity, respectively. The sample begins on March 6 and ends on April 9, 2020. Only trades that are greater than or equal to \$1 million in par volume, i.e. large trades, are included. Bonds that change credit grade are excluded. Clustered standard errors at the day and bond levels are shown in parentheses. $^*p<.1$; $^{**}p<.05$; $^{***}p<.01$.

	Dependent variable:							
	Risky-principal				Agency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SMCCF × Eligible	-24.22^{***}	-22.17^{**}	-29.02^{***}	-30.56^{***}	-6.45^{***}	-7.96^{***}	-1.81	-2.08
	(9.08)	(8.78)	(8.94)	(8.90)	(2.49)	(2.85)	(3.19)	(3.19)
SMCCF	6.70	4.25	5.12	6.66	4.43	5.76^{*}	-0.83	-0.63
	(9.27)	(10.06)	(9.91)	(9.96)	(2.82)	(3.03)	(3.16)	(3.18)
Eligible	-0.16	-9.19			-19.98^{***}	-7.83^{***}		
	(7.51)	(10.00)			(1.57)	(2.75)		
log(Amt outstanding)	-19.14^{***}	-26.76^{***}			-1.10	0.31		
	(2.67)	(3.17)			(0.98)	(0.94)		
log(Time-to-maturity)	19.55^{***}	22.15^{***}			2.49^{***}	4.99^{***}		
	(2.33)	(2.99)			(0.76)	(0.90)		
log(Age)	14.35^{***}	15.53^{***}			0.42	0.81		
	(2.43)	(2.99)			(1.28)	(1.28)		
Industry FE	Yes	Yes	No	No	Yes	Yes	No	No
Bond FE	No	No	Yes	Yes	No	No	Yes	Yes
Credit rating FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	29,941	28,992	29,941	28,992	8,983	8,367	8,985	8,367
Adjusted R^2	.02	.02	.02	.02	.13	.18	.28	.28

for transaction services. The demand equation is:

$$\log(x_{ht}/x_{lt}) = 1/\sigma \left[\theta_\tau + \varepsilon_t - (p_{ht} - p_{lt})\right],$$

where, with some abuse of notation, θ_τ is the average relative demand shifter in period $\tau \in \{A, B\}$, and ε_t is a mean zero shock. The supply equation is assumed to have the following form

$$\log(x_{ht}/x_{lt}) = b_0 + b_1 (p_{ht} - p_{lt}) + b_2 \log(N_t) + \eta_t, \tag{2.D.1}$$

where η_t is a mean zero supply shock, and N_t is the customer-to-dealer volume. Consistent with our model, N_t enters in the supply equation because it increases the marginal cost of providing transaction services. It does not enter the demand equation because the total utility for transaction services is linearly homogeneous: the utility of a given bundle of transaction service is the same for each dollar of transaction. We assume, moreover, that volume is smaller in the earlier period (A) than in the later period (B).[38] Formally, this can be written $\log(N_t) = \log(\overline{N}_\tau) + \xi_t$, where ξ_t is a mean zero shock, $\tau \in \{A, B\}$ and $\overline{N}_A > \overline{N}_B$.

Solving the system of simultaneous equations gives

$$p_{ht} - p_{lt} = \frac{\theta_\tau + \varepsilon_t - \sigma \left(b_0 + b_2 \log(\overline{N}_\tau) + b_2 \xi_t + \eta_t\right)}{1 + b_1 \sigma},$$

and

$$\log(x_{ht}/x_{lt}) = \frac{b_0 + b_2 \log(\overline{N}_\tau) + b_2 \xi_t + \eta_t}{1 + b_1 \sigma} + \frac{b_1 (\theta_\tau + \varepsilon_t)}{1 + b_1 \sigma}$$

[38] Using an unpaired two-sample t-test, we test whether that the volume is smaller in the earlier period (A) than in the later one (B). We first check whether the variance of log volume is similar in the two periods using an F-test. The null hypothesis is that the variance of log volume is the same in both period. The value of the test statistic is 0.32, corresponding to a p-value of 0.012. Hence, we fail to reject the equality of variances for log volume (in level variances seem different, which is why we did it in logs). Next, we do a one-sided t-test. The null hypothesis is that volume is smaller in the period A than in period B. The value of the test statistic is -0.97 corresponding to a p-value of 0.83. Again we fail to reject the null that volume is smaller in period A than in period B.

One sees that:

$$\mathbb{E}\left[p_{ht} - p_{lt} \mid t \in \tau\right] = \frac{\theta_\tau - \sigma\left(b_0 + b_2 \log(\overline{N}_\tau)\right)}{1 + b_1\sigma},$$

while

$$\mathbb{E}\left[\log(x_{ht}/x_{lt}) \mid t \in \tau\right] = \frac{b_1\theta_\tau + \left(b_0 + b_2 \log(\overline{N}_\theta)\right)}{1 + b_1\sigma}$$

It thus follows that, if $\theta_A = \theta_B$:

$$\sigma = -\frac{\mathbb{E}\left[p_{ht} - p_{lt} \mid t \in B\right] - \mathbb{E}\left[p_{ht} - p_{lt} \mid t \in A\right]}{\mathbb{E}\left[\log(x_{ht}/x_{lt}) \mid t \in B\right] - \mathbb{E}\left[\log(x_{ht}/x_{lt}) \mid t \in A\right]}.$$

Thus, an estimator of σ is obtained by replacing the conditional expectations by sample averages. To be more precise, let

$$Y_t = \begin{pmatrix} p_{ht} - p_{lt} & \log(x_{ht}/x_{lt}) \end{pmatrix}'$$

denote the vector of observations at time t. Assume for now that all the vector of disturbances, $(u_t, v_t, w_t)'$ are IID over time with finite covariance matrices. The estimate of the mean vector over periods A and B are:

$$\hat{Y}_\tau = \frac{1}{T_\tau} \sum_{t \in \tau} Y_t$$

where T_τ denotes the number of observations in period τ. Then, the weak Law of Large Numbers implies that sample means converge in probability to their population counterpart, $\overline{Y}_\tau$, as T_τ goes to infinity. The Central Limit Theorem implies that $\sqrt{T_\tau}\left(\hat{Y}_\tau - \overline{Y}_\tau\right)$ is asymptotically normally distributed with mean zero and covariance matrix S. Moreover, these two random variables are also independent. Since

$\left(Y_t - \overline{Y}_\tau\right)$, $\tau \in \{A, B\}$ are IID, an unbiased and consistent estimator of S is

$$\hat{S} = \frac{T_A - 1}{T_A + T_B - 2} \frac{1}{T_A - 1} \sum_{t \in A} \left(Y_t - \hat{Y}_A\right) \left(Y_t - \hat{Y}_A\right)' + \frac{T_B - 1}{T_A + T_B - 2} \frac{1}{T_B - 1} \sum_{t \in B} \left(Y_t - \hat{Y}_B\right) \left(Y_t - \hat{Y}_B\right)'$$

$$= \frac{1}{T - 2} \left(\sum_{t \in A} \left(Y_t - \hat{Y}_A\right) \left(Y_t - \hat{Y}_A\right)' + \sum_{t \in B} \left(Y_t - \hat{Y}_B\right) \left(Y_t - \hat{Y}_B\right)' \right)$$

The estimate of σ can be written as

$$\hat{\sigma} = f\left(\hat{Y}_A, \hat{Y}_B\right) = -\frac{\hat{Y}_{A2} - \hat{Y}_{A1}}{\hat{Y}_{B2} - \hat{Y}_{B1}},$$

where the 1 and 2 subscripts denote the first and second coordinates of the Y vector. By standard delta method, we thus have that

$$\begin{aligned} \sqrt{T}\left(\hat{\sigma} - \sigma\right) &= \sqrt{T}\left(f\left(\hat{Y}_A, \hat{Y}_B\right) - f\left(\overline{Y}_A, \overline{Y}_B\right)\right) \\ &= \sqrt{T}\left(\frac{\partial f}{\partial Y_A'}\left(\hat{Y}_A - \overline{Y}_A\right) + \frac{\partial f}{\partial Y_B'}\left(\hat{Y}_B - \overline{Y}_B\right)\right) \\ &= \sqrt{\frac{T}{T_A}}\sqrt{T_A}\frac{\partial f}{\partial Y_A'}\left(\hat{Y}_A - \overline{Y}_A\right) + \sqrt{\frac{T}{T_B}}\sqrt{T_B}\frac{\partial f}{\partial Y_B'}\left(\hat{Y}_B - \overline{Y}_B\right). \end{aligned}$$

By symmetry it is clear that $\sqrt{T_A}\frac{\partial f}{\partial Y_A'}\left(\hat{Y}_A - \overline{Y}_A\right)$ and $\sqrt{T_B}\frac{\partial f}{\partial Y_B'}\left(\hat{Y}_B - \overline{Y}_B\right)$ have the same asymptotically normal distribution, with mean 0 and variance

$$\frac{\partial f}{\partial Y_A'} S \frac{\partial f}{\partial Y_A}, \text{ where } \frac{\partial f}{\partial Y_A'} \equiv \begin{pmatrix} \frac{1}{\overline{Y}_{B2} - \overline{Y}_{A2}} & -\frac{\overline{Y}_{B1} - \overline{Y}_{A1}}{\left(\overline{Y}_{B2} - \overline{Y}_{A2}\right)^2} \end{pmatrix}.$$

It thus follows that the estimate $\hat{\sigma}$ is asymptotically normal with mean σ and standard deviation

$$\left(\frac{1}{T_A}+\frac{1}{T_B}\right)^{1/2}\left(\frac{\partial f}{\partial Y_A'}S\frac{\partial f}{\partial Y_A}\right)^{1/2}$$

A consistent estimate of the standard deviation is found by replacing S by $\hat{S}$, and the population means in $\partial f/\partial Y_A'$ by their sample counterparts.

For the estimation we set the early period A to run between 2020-01-15 and 2020-02-14, and the late period B to run between 2020-06-01 and 2020-06-30. We use the raw series for x_h and x_l, not their moving average, so as not to artificially reduce standard errors. We obtain an estimate of $\hat{\sigma} = 100.09$ with a standard deviation of 15.4.

2.D.2 A second binary IV

One may argue that the assumption underlying our IV estimation, $\mathbb{E}_{\text{Jan}}[\theta_t] = \mathbb{E}_{\text{Jun}}[\theta_t]$, is too strong. A key concern is that, by May 2020, θ_t has not returned to its pre-crisis level. Instead, it has stabilized to a “new normal”: it is lower than at the height of the crisis, but higher than before the crisis.

To address this concern, we re-estimate σ under this “new normal” assumption. We take the initial period T_A to be from May 1 to May 15, 2020 and the final period T_B to be from June 15 to June 30, 2020. Our assumption is that the average level of θ_t is the same in both periods. Figure 2.4.2 suggests that the volume of customer-to-dealer trades declined significantly between periods T_A and T_B: following the same logic as in Section 2.4.3.1, we argue that this decline in volume shifted the supply curve for risky-principal trade, leading to the estimate

$$\hat{\sigma} = -\frac{\frac{1}{T_B}\sum_{t\,:\,Z_t=1}(p_{ht}-p_{lt}) - \frac{1}{T_A}\sum_{t\,:\,Z_t=0}(p_{ht}-p_{lt})}{\frac{1}{T_B}\sum_{t\,:\,Z_t=1}\log(x_{ht}/x_{lt}) - \frac{1}{T_A}\sum_{t\,:\,Z_t=0}\log(x_{ht}/x_{lt})},$$

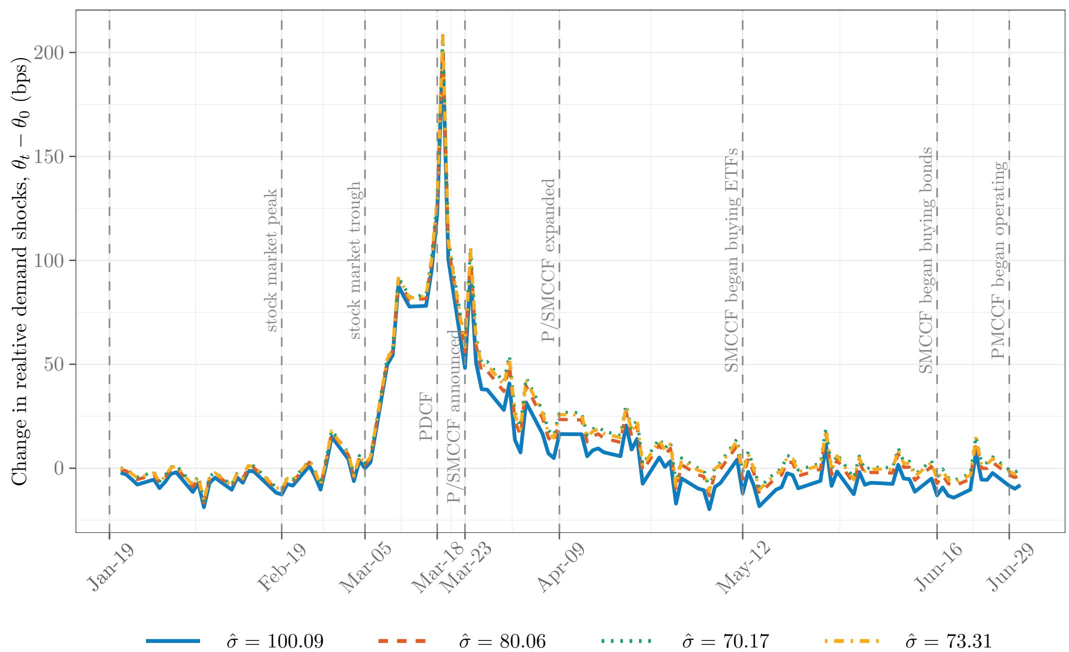
This calculation leads to an estimate of $\hat{\sigma} = 80.06$ which is close to our other estimates. Given this value of the semi-elasticity of demand, we can infer a new series for the demand shock as we did in Section 2.4.4.

$$\theta_t - \theta_0 = \hat{\sigma}\left[\log\left(x_{lt}/x_{ht}\right) - \log\left(x_{l0}/x_{h0}\right)\right] - \left(p_{ht} - p_{lt} - \left(p_{h0} - p_{l0}\right)\right).$$

As before, we take $t = 0$ to be January 2, 2020.

In Figure 2.D.1, we plot the implied demand shocks from the four different estimates of the semi-elasticity parameter σ: two binary IVs (solid blue and dashed red line) and two high-frequency IVs (green dotted and orange dashed-dotted lines). As we can see from Figure 2.D.1, the demand shock is at a stable level between early May and late June. Notice that the "new normal" assumption only implies that the average level in those sub-periods would be the same and does not imply stability in between those periods. More importantly, we do not make an assumption on the level at which the demand shock would stabilize, yet we find that it is remarkably close to the pre-crisis level.

The figure reveals that our estimates are quite robust: despite the fact that these estimation procedures derive from significantly different identification assumptions and strategies, the implications for our estimate of θ_t—and hence the discussion in Section 2.4.4—is essentially unchanged.



2.D.1. The change in the estimated relative demand shocks for risky-principal trades for different estimates of the semi-elasticity parameter $\boldsymbol{\sigma}$. This figure plots the time-series of the change in the estimated relative demand shocks for risky-principal trades relative to a pre-crisis benchmark on January 2, 2020, $\theta_t - \theta_0$, implied from the logit demand, $p_{ht} - p_{lt} = -\sigma \log(x_{ht}/x_{lt}) + \theta_t$, for four values for parameter σ.

2.D.3 High-frequency IVs

In this section we provide additional details and robustness results for the high-frequency IV method discussed in Section 2.4.3.1. Table 2.D.1 presents our estimates for σ from an overidentified IV. All specifications contain a post-crisis dummy variable allowing for a one time shift in the demand curve. This method gives us an estimate of σ of 73.10, that lies between the two individual IV estimates in columns (1) and (2) from Table 2.4.1. The weak instrument test statistic is 11.172. Since here we have an overidentified system, we obtain the Sargan's J test statistics. The value of the test statistic is 0.057 and the p-value of 0.81. So we fail to reject the

2.D.1. Estimating the logit demand parameter σ: overidentified IV. This table presents the IV estimate of the logit demand parameter σ. Both log volume and number of customer-to-dealer trades are used as instruments. The pre-crisis runs from January 3, 2020 until February 29, 2020. The post-crisis data begins on April 15, 2020 and runs until July 31, 2020. We exclude holidays, weekends and half trading days. Standard error are the maximum of robust and the usual standard errors. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

	Dependent variable:
	$(p_h - p_l)$ IV (num & vol)
	(1)
$\log(x_h/x_l)$	73.10**
	(31.04)
Post-crisis	7.78
	(5.16)
Constant	77.97***
	(25.81)
Observations	113
Adjusted R^2	.39

validity of the overidentification restrictions.

In Table 2.D.2, we present the first stage of our IV regressions for each instrument. We find that both log number and volume of trades are valid instruments for $\log(x_h/x_l)$. The F-stat for the weak instruments test is 16.731 in the case of volume and 22.509 for the number of trades. In both cases we therefore reject the null that the instruments are weak at the 1% level. Both instruments have a correlation of 0.88.

2.D.2. The first stage of the IV regressions for estimating the logit demand parameter $\boldsymbol{\sigma}$. We regress the log ratio of fraction of risky-principal and agency trades on the log of the seasonally-adjusted daily aggregate number and volume of trades in columns (1) and (2), respectively. The pre-crisis runs from January 3, 2020 until February 29, 2020. The post-crisis data begins on April 15, 2020 and runs until July 31, 2020. We exclude holidays, weekends and half trading days. Standard errors are given by the maximum of robust and the usual standard errors. *p <.1; **p <.05; ***p <.01.

	Dependent variable:	
	$\log(x_h/x_l)$	
	(1)	(2)
log(Volume of trades)	-0.22^{***}	
	(0.05)	
log(Number of trades)		-0.44^{***}
		(0.09)
Post-crisis	-0.12^{***}	-0.13^{***}
	(0.02)	(0.02)
Constant	0.82^{***}	0.84^{***}
	(0.01)	(0.01)
Observations	113	113
Adjusted R^2	.45	.47

2.E Relative Preference Shocks for Eligible/Ineligible Bonds

To explore the relationship between customers' preferences and the Fed's interventions in greater depth, we can estimate the time path of θ_t separately for those bonds that were eligible vs. ineligible for purchase according to the March 23 announcement of the corporate credit facilities.

According to the eligibility requirements specified in the March 23, 2020 announcement of the SMCCF, we divide the whole sample of transactions into two sub-samples, one for eligible bonds and the other for ineligible bonds. We identify eligible bonds as the ones with an investment-grade rating on March 23, a time-to-maturity of five years or less by March 23, and its issuer domiciled in the U.S. In particular, we calculate each bond's time-to-maturity by March 23 as the gap in days between its maturity date and the date of March 23. We set all the other bonds as ineligible ones. Then we use the two sub-samples of transactions to correspondingly generate two samples of observations on prices and relative quantities, and separately estimate the time path of θ_t for each sample.

Then given the estimate of the semi-elasticity parameter σ for all bonds, using the time series of the price premium and the ratio of the risky-principal trades for eligible and ineligible bonds, we can infer the sequence of shocks to customers' relative demand shock for eligible and ineligible risky-principal trades:

$$\theta_t^j - \theta_0^j = p_{ht}^j - p_{lt}^j - \left(p_{h0}^j - p_{l0}^j\right) - \sigma\left[\log\left(x_{lt}^j / x_{ht}^j\right) - \log\left(x_{l0}^j / x_{h0}^j\right)\right],$$
$$j \in \{\text{eligible}, \text{ineligible}\}.$$

Using the estimate $\hat{\sigma} = 100.09$ from our binary IV approach, Figure 2.4.5 plots the time-series of the change in θ relative to a pre-crisis benchmark (on January 2, 2020) for eligible (solid blue line) and ineligible bonds (dashed red line).

CHAPTER 3

Trading Relationship in the US Corporate Bond Market

This is joint work with Diego Zúñiga.

"We're greedy, but long-term greedy, not short-term greedy."

Gus Levy, Goldman Sachs Senior Partner 1969-1976.

3.1 Introduction

The US corporate bond market is organized as a decentralized over-the-counter (OTC) market where customers and dealers trade. In this market when a dealer wishes to initiate a trade with another dealer, they typically contact individual dealers and then bargain over the terms of trade. [1] Inter-dealer transactions are typically modeled as either static trading networks or as using a technology that is independent of the past trades of the dealer, Duffie et al. (2007).

[1]While request for quote trading venues allow participants to contact multiple counterparties simultaneously they represent only a small fraction of the volume traded, O'Hara and Alex Zhou (2021).

However, one can also view these interactions as economically meaningful relationships. A dealer pair has a trading relationship if they repeatedly trade over time. Such a trading relationship is economically meaningful if trading with the same small number of counterparties affects the terms of trade that the dealers bargain over or confers other value to the dealer. This may include the ability to share risks with other dealers, access market information, or reduced inter-dealer transaction costs.

Our paper proceeds by first documenting the patterns of inter-dealer trading that we observe in the US corporate bond market. We use the academic enhanced version of the Trade reporting and compliance engine (TRACE) data set from the Financial Industry Regulatory Authority (FINRA) which contains anonymized dealer identities allowing us to derive several stylized facts about the trading patterns that exist in the corporate bond market. As is true in many OTC financial markets the US corporate bond market resembles a core-periphery network; a small number of highly connected dealers as well as a much larger number of peripheral dealers who trade with only a small number of other dealers.

We show that the typical dealer trades with only a small number of other dealers. Each month the median dealer trades with only three other dealers (mean 15 dealers per month). This is despite the fact the median dealer transacts in the inter-dealer market 22 times per month (mean is 592 trades a month). Moreover, the group of dealers whom a dealer trades with is highly persistent over time. The unconditional probability that a dealer trades within a 30-day period with another dealer is less than 0.02 whereas conditional on trading in the previous 30 days the probability of trading in the next 30 days is 0.42. The probability is increasing in the number of 30-day periods of continuous trading. For dealer pairs that have traded for over two years continuously the probability that they continue to trade is over 0.95. We argue that this is because dealers trade with a small group of other dealers whom they have a relationship with.

We find the longer a dealer pair have been continuously trading the larger the volume the pair trade. One additional month of continuous trading increases the par volume traded between the pair by over $43,000,000. In our view this reflects the fact that dealers gradually learn which other dealers are "good" counterparties which leads them to further concentrate their trading volume with these dealers. Here we envisage a process that is similar to how firms can acquire a reputation for quality.

These facts lead us to construct various measures of relationship strength such as the number of continuous months a dealer pair has been trading, the share of the volume they intermediate with a dealer, and whether the dealer pair traded in the previous 30 days. We show that these measures are highly correlated. To measure inter-dealer trade costs, we construct the imputed round-trip cost (IRC) for riskless-principal trades. Using triple interacted bond, time, and dealer fixed effects to account for bond, time, and dealer heterogeneity we regress the dealer's trade costs on measures of relationship strength. Our regressions reveal that trading costs are lower when a dealer transacts with a counterparty whom they have a strong trading relationship.

Motivated by the stylized facts we build a model of trading in the US corporate bond market. The model is set in discrete time and features both dealers and customers. Dealers are long-lived agents and customers exist for one period only. At the beginning of the period with a certain probability each dealer meets with a potential customer. Customers enter the market with a desire to trade and meet one and only one dealer. After meeting with a dealer, the customer and dealer bargain over the price and quantity that they wish to trade. The pair may decide to trade any fraction of the customer's initial position. However, the marginal cost for the dealer of trading with the customer is increasing in volume.

After meeting with a customer and committing to trade a particular volume, the dealer may then enter the inter-dealer market. The dealer may contact any dealer without a customer and transact with them. Since the cost of intermediating customer volume is convex in volume there are gains to trade. After contacting another dealer, they bargain over the terms of trade, with both parties having some market power. After the dealers agree to terms of trade all trades are settled and the period ends.

In our model market power in the inter-dealer market leads to double marginalization. Double marginalization causes the dealer who locates the customer to intermediate a sub-optimal amount, reducing aggregate surplus. However, since dealers repeatedly interact with each other, they may develop a reputation or relationship with another dealer. We derive conditions under which such relationships can exists and when they restore optimality.

Finally, we use our model to derive predictions about the impact of relationships on the trading patterns in across different bond markets. Since dealers tend to specialize in trading only a subset of all corporate bonds, some bond markets tend to have strong relationships while in other markets relationships will be weaker. We show that customer trading volume is lower when dealers in the bond market have weaker relationships. Intuitively, it is more costly for dealers to trade and so transaction costs will be higher for customers. Customers therefore choose to trade a smaller volume with dealers.

The remainder of the paper is organized as follows. Section 3.2 provides an overview of the relevant literature. Section 3.3 describes the data and the process we used to clean it. Section 3.4 describes some facts about relationships. Section 3.5 provides evidence of how relationship impact inter-dealer transaction costs. Section 3.6 presents the model. Section 3.7 examines how customer transaction volumes are affected by trading relationships. Section 3.8 concludes.

3.2 Related Literature

While this paper makes several distinct contributions to the literature, we highlight several aspects of this study that are particularly important. This paper contributes to the literature on over-the-counter (OTC) financial markets, particularly the literature that studies the US corporate bond market, Duffie et al. (2005), Duffie et al. (2007), Atkeson et al. (2015), and Hugonnier et al. (2019). In contrast to these papers our contribution is to explicitly allow for economically meaningful relationships in the inter-dealer market. In our model the existence of market power in the inter-dealer market leads to double marginalization. However, the repeated nature of the interactions allows dealers to build relationships which overcome this friction.

Most closely related to our paper are Afonso et al. (2014), Di Maggio et al. (2017) and Hendershott et al. (Forthcoming) which study relationships among different parties in various OTC financial markets. Hendershott et al. (Forthcoming) studies relationships between corporate bond dealers and insurers, while Di Maggio et al. (2017) uses the exit of a large dealer during the 2008 financial crisis to empirically analyze the value of relationships that other trading counterparties had with the dealer. Relative to these papers our contribution is in documenting additional stylized facts about inter-dealer relationships, such as the dynamics of these relationships in the data and to provide an explicit model for inter-dealer relationships.

Finally, our paper is related to the literature on underlying trading structure of OTC markets, An (2019), Eisfeldt et al. (2018), Bao et al. (2018), O'Hara and Alex Zhou (2021), Colliard et al. (2018), Babus and Kondor (2018), Bessembinder et al. (2018), Goldstein and Hotchkiss (2020b), Li and Schürhoff (2019), Farboodi (2014), Babus and Hu (2018). Our contribution to this literature is to document the formation of long-term repeated interactions among dealer pairs. We show the probability of continuing to trade depends on the length of time a dealer pair have been trading and that the volume a dealer pair trades also depends on how long the pair have been trading.

3.3 Data and Cleaning

Our paper uses two data sets, corporate bond transaction level data from the academic enhanced version of the TRACE data base from FINRA and Mergent Fixed Income Securities Database (FISD). The TRACE database contains the universe of trade reports from US corporate bond dealers and our version runs from June 2002 until the end of 2016. Unlike the regular version the academic enhanced version of TRACE contains masked dealer identities and uncapped trade volumes. While customer identities remain unknown, for inter-dealer trades we can see which market participants trade with each other over time. The Mergent FISD data set provides us with bond characteristics, such as issue and maturity date, and credit rating. In total we have 100,442,441 unique trades and 44,746 bonds. Table 3.3.1 provides further summary statistics of the data.

3.3.1. This table contains summary statistics for our data. Each unit of observation represents a trade. Volume is in US dollars and represents transaction volume. Price is as a percent of par. Our sample runs from June 2002 until December 2016.

Statistic	Mean	Median	St. Dev.
Price	100.21	101.18	15.42
Volume	615,899.50	25,000	2,145,388.00

We clean the TRACE data using the standard procedure in Dick-Nielsen (2014). Additionally, we remove bonds that are not in the FISD data bases. We remove transactions with erroneous transaction dates, i.e., those that take place before the issue date or after the maturity date, with a volume larger than the issue size. To avoid the influence of very large outliers we truncate the trade volume at the 99.99% percentile. We also remove trades that don't take place during a normal business day, or with a negative reported price. We also appropriately correct for give-up and locked-in-trades to ensure that we have the actual parties who are transacting. We delete duplicate trade reports where appropriate.

Furthermore, as is standard, we remove unusual bonds and bonds that are infrequently traded. Unusual bonds include securities with an offering amount of less than $1 million. We define infrequently traded as bonds that have fewer than 50 trades in total, are traded for fewer than 3 dealers or are in TRACE for fewer than 60 days, or with transaction on less than 10 days in total. Finally, we implement a price error filter in the spirit of Rossi (2014) winzorising prices that are very far from the left and right rolling median transaction price for the bond [2].

In total our data set has 3685 unique dealers. The mean number of dealers who

[2] See appendix for further details

trade in a month is 1120 dealers (median 1110, standard deviation 152.95). In table 3.3.2 we have dealer month summary statistics. The median dealer trades with only 4 counterparties per month, with 20 dealers being the mean number of non-customer counterparties.

In table 3.3.3 we furthermore see that a dealer pairs that trade with each other typically trade many times in a given month. We find that a dealer pair who trade at least once in a month trade a mean of 42 times (median 3 trades). Suggesting that dealers tend to concentrate their trades rather than spread them among all the possible counterparties.

3.3.2. This table contains summary statistics for a dealers monthly trades. Each unit of observation represents a dealer-month. Volume is millions of US dollars and represents transaction volume. Which is defined as par-volume times price as percentage of par divided by 100. Customers are defined as trades with entities without a FINRA ID. Our sample runs from June 2002 until December 2016.

Statistic	Mean	Median	St. Dev.
Customer volume	315.09	0.97	2,287.29
Interdealer volume	460.37	2.09	2,848.07
Number of contra-parties	19.95	4	41.66
Bonds traded	152.87	12	530.63
Number of trades	838.21	31	4,029.79

3.4 Facts about Relationships

The summary statistics in tables 3.3.2 and 3.3.3 suggest that dealers are concentrating their trades with just a small number of the available dealers. Figure 3.4.1 illustrates the trading patterns of the US corporate bond market. In figure 3.4.1 the

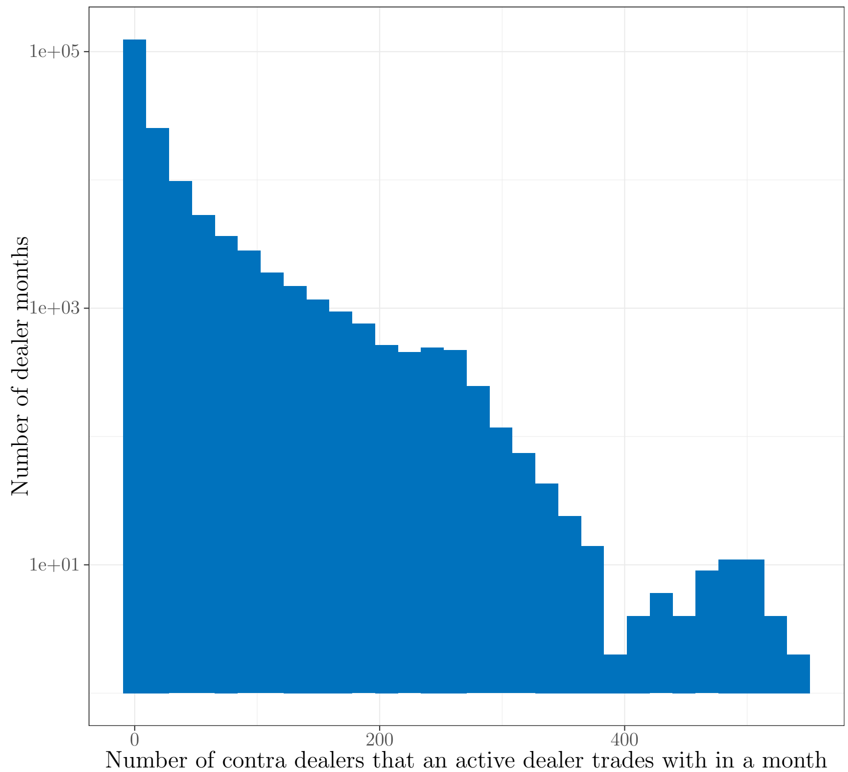
3.3.3. This table contains summary statistics for a dealers pair in a given monthly. Each unit of observation represents a dealer-pair-month. Volume is millions of US dollars and represents transaction volume. Our sample runs from June 2002 until December 2016.

Statistic	Mean	Median	St. Dev.
Bilateral volume	8.43	0.39	52.11
Number of bilateral trades	23.26	3	176.87
Proportion of inter-dealer from this pair	0.05	0.004	0.15

dark region represents the part of the inter-dealer network where at a pair of dealers traded with each other at least once between June 2002 and the end of 2016. The light-grey region represents pairs of dealers who did not trade. As is common in many financial markets, see Eisfeldt et al. (2018) for example, the network exhibits a core-periphery structure. The vast majority of dealers never trade with each other. Most dealers only ever trade with a small number of dealers. A small subset of dealers (the core) trade with many and almost all dealers trade with at least one member of the core. Our paper seeks to interpret this structure through the lens of inter-dealer relationships. We first illustrate some stylized facts about the trading structure in the US corporate bond market.

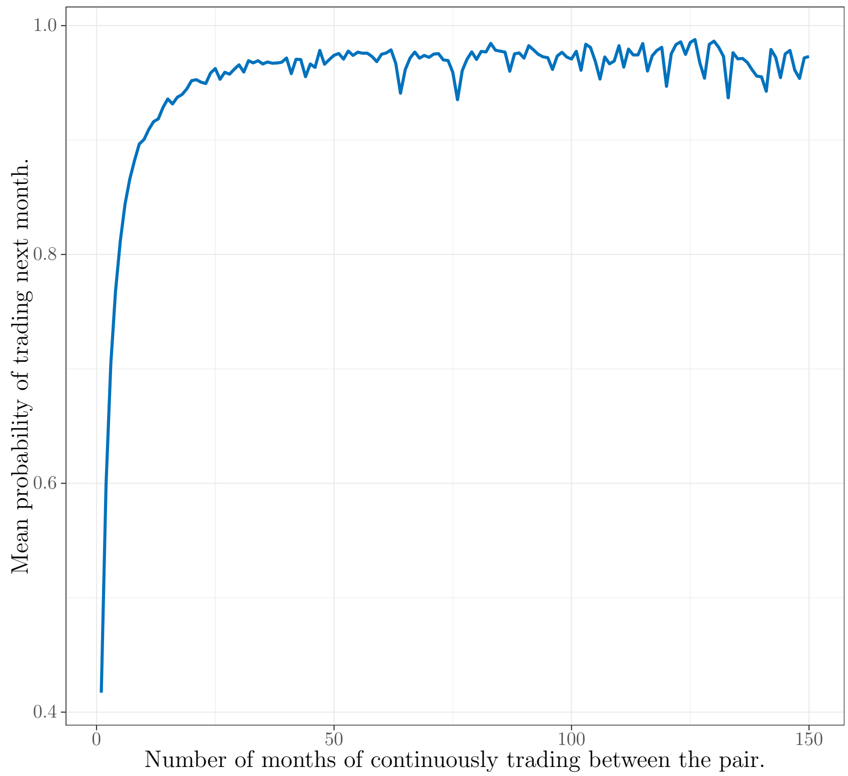
In figure 3.4.2 we plot the distribution of the number of dealer counterparties active dealers have in a given month. An active dealer is a dealer who trades at least once in that month, with either another dealer or an end customer. Thus, trading with only customers in a given month mean that the dealer has no contra dealers. This plot demonstrates that most dealers trade with only a small number of other dealers out the thousands that are active in a given month. We also see that a distinct core of dealers who have significantly more counterparties than the others in the market.

3.4.1. This figure plots the aggregate US corporate bond market. We define a pair of dealers to be connected if they ever trade with each other. Dark areas represent parts of the network that are connected and the grey region represents the unconnected dealers. We order dealers by the number of connections they have.



3.4.2. This figure plots the number counterparties a dealer has in a given month. A unit of observation is a dealer month. Only dealers who trade at least once (either with another dealer or with a customer) in the given calendar month are included in the plot. Customers are defined as trades with entities without a FINRA ID and are not counted as a trading party due our inability to identify them. Our sample runs from June 2002 until December 2016.

The first stylized fact that we document is that the probability that a dealer pair trade increases the longer two dealers have been trading. We define the number of months two dealers have been continuously trading in month m as the maximum number of calendar months a pair of dealers have been trading in months prior to m. Thus, if a pair of dealers were trading in January and February then in March this variable will equal 2. If however the dealers were not trading in February then this variable will equal zero.



3.4.3. This figure plots the number of months a pair of dealers have been trading against the mean probability that they will continue to trade. We see an increasing relationship between the number of months a pair of dealers have been trading and the likelihood that they will continue to trade.

In figure 3.4.3 we plot the mean probability that a dealer pair will continue to trade conditional on them trading continuously for 1, 2, 3, ... months. We cap the number of months at 150 due to the limited span of our data. As the number of months a pair of dealers have been continuously trading increases, the likelihood that they continue to trade increases. However, after 24 months the probability reaches about 0.95, and from there on the probability remains approximately constant.

A more formal analysis of figure 3.4.3 is in table 3.4.2. Here we regress the probability of trading next month on the number of months the pair of dealers have been continuously trading. We include a variety of fixed effects, including month, dealer, and contra dealer fixed effects. These fixed effects remove variation due to time-invariant dealer and contra covariates including size, market position, and matching ability, as well as variation due to changes that occur to the market structure over time. All standard errors are double clustered at the month and dealer level. We find that an extra month of trading increases the probability that a dealer pair continue to trade by between 0.0034 and 0.0056 depending on the specification.

The results of table 3.4.1 indicate that dealers form trading partnerships with each other. The likelihood that a pair of dealers continue to trade depends on their history of bilateral trading. Dealers who traded many more frequently in the past are more likely to continue to trade compared to those that traded infrequently. This is at odds with the standard random search model used frequently used to describe OTC markets, Hugonnier et al. (2019) and Üslü (2019), where the probability of matching should not depend on the history of trades, conditional on the agents type.

Another channel through which relationships may form is by the trading volume between agents. It may be that dealers who have stronger relationships tend to trade a larger total volume compared to those with weaker relationships. In figure 3.4.4 we plot the mean trading volume in a given month, across all pairs of dealers who have been trading continuously for 1,2,3,... months. As was the case for figure 3.4.3 we exclude observations with more than 150 months of continuously trading due to the limited span of our data set. The red line is the line of best fit fitted using least squares.

From figure 3.4.4 we can see that on average as the number of months a pair of dealers have been trading continuously increases the total trading volume between

3.4.1. This table contains the results of the regression of the probability that a dealer pair will continue to trade on the number of months that they have been continuously trading. Each unit of observation represents a dealer-pair-month. The number of months that a pair have been contagiously trading is defined in the text. We exclude all trades by non-FINRA members. Our sample runs from June 2002 until December 2016. Standard errors are double clustered at the dealer and month level.

Dependent Variable:	Trade next month				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
(Intercept)	0.6408***				
	(0.0058)				
Months continuous trade	0.0054***	0.0046***	0.0047***	0.0056***	0.0034***
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Fixed-effects					
Dealer	No	Yes	No	No	Yes
Contra dealer	No	No	Yes	No	Yes
Time (monthly)	No	No	No	Yes	Yes
Fit statistics					
Observations	3,416,256	3,416,256	3,416,256	3,416,256	3,416,256
R^2	0.08701	0.12386	0.12106	0.1086	0.20434
Within R^2	–	0.06108	0.06246	0.0919	0.02792

Two-way (Time (monthly) & Dealer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

them increases. Dealers who have been continuously trading for 100 months (12 years) trade about $40 million on average compared to a just over $20 million on average for dealers who have been trading for 50 months.

In the same vain as table 3.4.1 in table 3.4.2 we regress the total trading volume between a pair of dealers on the number of months that they have been continuously trading as well as dealer, contra, and date (month level) fixed effects. Again, standard errors are double clustered at the time (month) and dealer level. We find that every extra month a dealer pair have been trading the total monthly volume they trade increases by between $300,000 and $433,000 depending on the specification.

Taken together these facts are indicate that dealers tend to concentrate their trading volume with a small number of counterparties, rather than spreading it amongst the entire dealer network. Furthermore, given that the probability of trading with a dealer is increasing in the number of months that the pair have been trading this also suggests that dealers tend to continue to trade with the *same* small number of counterparties. This suggests that dealers form long term relationships with each other and raises the question as to what the economic motivation is for forming such long term relationships is.

3.4.1 Measuring Relationship

We first discuss several methods of measuring relationships between dealers. Here we draw on the existing literature that's related to relationship in OTC markets such as Hendershott et al. (Forthcoming), Di Maggio et al. (2017) and Afonso et al. (2014), as well as proposing our own measures. A first simple measure of whether a pair have a trading relationship is to define a pair to have a relationship if they traded

3.4.4. This figure plots the number of months a pair of dealers have been trading against the mean probability that they will continue to trade. We see an increasing relationship between the number of months a pair of dealers have been trading and the likelihood that they will continue to trade.

with each other in the previous 30 days. This simple measure has the advantage of being invariant to the scale of the dealer. However, it fails to differentiate between relationships in terms of importance to the dealer.

Another measure one could use to measure relationships is the number of months (defined as 30 day periods rather than calendar months) a pair of dealers have been trading with each other. This measure allows us to differentiate between relationships that are more and less important to the dealer. This measure however suffers from the fact that the span of our data is limited. Furthermore, it implicitly assumes that

3.4.2. This table contains the results of the regression of the total trading volume between pair of dealers on the number of months that they have been continuously trading. Each unit of observation represents a dealer-pair-month. The total trading volume is the value of the trades that the pair of dealers execute with each other measured in millions of US dollars. The number of months that a pair have been contagiously trading is defined in the text. We exclude all trades by non-FINRA members. Our sample runs from June 2002 until December 2016. Standard errors are double clustered at the dealer and month level.

Dependent Variable:	Volume (millions USD)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
(Intercept)	2.226^{***}				
	(0.3407)				
Months continuous trade	0.4064^{***}	0.3564^{***}	0.3576^{***}	0.4329^{***}	0.2998^{***}
	(0.0656)	(0.0522)	(0.0592)	(0.0708)	(0.0451)
Fixed-effects					
Dealer	No	Yes	No	No	Yes
Contra dealer	No	No	Yes	No	Yes
Time (monthly)	No	No	No	Yes	Yes
Fit statistics					
Observations	3,416,256	3,416,256	3,416,256	3,416,256	3,416,256
R^2	0.03628	0.12422	0.11986	0.04003	0.20873
Within R^2	–	0.02759	0.02773	0.0392	0.01624

Two-way (Time (monthly) & Dealer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

strong relationships take time to form. This may not be the case if for example a dealer enters the market and immediately wishes to form strong relationships with others.

A way to overcome the issue that dealers who have been in the data longer have stronger relationships is to normalize the the number of months the two dealers have been trading by dividing by the time they have been active. Thus, we can a relationship measure as the number of 30 day periods a pair of dealers have been trading divided by the total number of 30 day periods they have been trading in TRACE.

We can also measure relationships between pairs of dealers as a transformation of the total volume that they intermediate between the pair of dealers in the previous 30 days. Since the the distribution of volume between pairs of dealers is right skewed a log transformation seems most appropriate here. To account for the possibility of zeros we use a the volume + 1.

The final measure we propose is the proportion of total volume intermediated between the dealer in the previous 30 days. That is the total volume intermediated between dealers i and j in the previous 30 days divided by the total inter-dealer volume of dealer i [3]. This has the advantage of being bounded unlike the log of the volume + 1 and not subject to the entry and exit concerns like the number of months of continuous trading between the pairs.

In table 3.4.3 we illustrate some summary statistics for the proposed relationship measures. A unit of observation here is an inter-dealer trade and the associated relationship measures with that trade. To mitigate for exit and entry consideration we exclude the first 30 days a dealer is in the market. Table 3.4.3 makes clear that with the exception of whether the dealer's traded in the previous month there is considerable variation in these relationship measures. In table 3.4.4 we display the correlation matrix of these relationship measures. We can see that these measures are positively correlated with each other.

---

[3] We may divide by the the total inter-dealer volume of j. The choice of whether to divide by i or j is determined by which dealer is the reporting party of the transaction.

3.4.3. This table contains the summary statistics for the relationship measures outlined in the text. An observation is an inter-dealer trade and the associated relationship measures with that inter-dealer trade. The relationship measures are defined in the text. We exclude all trades by non-FINRA members and the relationship measures are only defined after a dealer has been in the sample for more than 30 days. Our sample runs from June 2002 until December 2016.

Statistic	Mean	Median	St. Dev.
Traded in previous 30 days	0.98	1.00	0.13
Number of 30 day periods trade	44.49	34.00	39.18
Proportion of 30 day periods trade	0.56	0.55	0.36
95^{th} percentile time gap	0.95	1.00	0.23
30 day rolling proportion of trades	0.09	0.02	0.19
Log 30 day rolling volume + 1	16.39	16.80	3.13

3.4.4. This table contains the correlation matrix for the relationship measures outlined in the text. An observation is an inter-dealer trade and the associated relationship measures with that inter-dealer trade. The relationship measures are defined in the text. We exclude all trades by non-FINRA members and the relationship measures are only defined after a dealer has been in the sample for more than 30 days. Our sample runs from June 2002 until December 2016.

Traded in previous 30 days	1	0.15	0.20	0.53	0.06	0.67
Number of 30 day periods trade	0.15	1	0.57	0.23	0.07	0.33
Proportion of 30 day periods trade	0.20	0.57	1	0.30	0.06	0.42
95^{th} percentile time gap	0.53	0.23	0.30	1	0.08	0.57
30 day rolling proportion of trades	0.06	0.07	0.06	0.08	1	0.30
Log 30 day rolling volume + 1	0.67	0.33	0.42	0.57	0.30	1

3.5 Transaction Costs

One way in which the relationships may lead to economic benefits for market participants is through a reduction in transaction costs. When a dealer trades with a counterparty whom they have a strong trading relationship with we would anticipate that they face lower transaction costs compared to a dealer whom they have a weaker trading relationship.

Measuring transaction costs for dealers is challenging due to the difficulty in inferring which dealer is demanding and which dealer is supplying liquidity. In TRACE both dealers must submit trade reports and, unlike customers, either could be the end buyer or seller.

Following Feldhütter (2012), one method for dealing with these difficulties is examine matched inter-dealer transactions. These are trades where dealer 1 trades with dealer 2 who immediately reverses that transaction with with dealer 3. The net effect in these trades is for dealer 1 to transact with dealer 3. For these trades since dealer 2 has prearranged a transaction for dealer 1 we can conclude that dealer 2 is providing liquidity to dealer 1. We can then measure the trade costs the spread between the buy and sell price of dealer 2, or

$$\text{IRC Spread}_{ijkst} = \frac{\text{Buy price of dealer } 2_{ijkst} - \text{Sell price of dealer } 2_{ijkst}}{2(\text{Buy price of dealer } 2_{ijkst} + \text{Sell price of dealer } 2_{ijkst})}, \quad (3.5.1)$$

Where the subscripts in equation 3.5.1 are read as i for dealer 1, j for dealer 2, k for dealer 3, s for security s and t for time.

We winzorize our measures of transaction costs at the 99.9% and 0.1% levels to avoid the influence of outliers on our results and multiply by 10,000 to convert to basis points. Table 3.5.1 contains summary statistics for our inter-dealer IRC measure.

To test whether stronger relationships lead to reduce trade costs we regress the

3.5.1. This table contains the correlation matrix for the relationship measures outlined in the text. An observation is an inter-dealer trade and the associated relationship measures with that inter-dealer trade. The relationship measures are defined in the text. We exclude all trades by non-FINRA members and the relationship measures are only defined after a dealer has been in the sample for more than 30 days. Our sample runs from June 2002 until December 2016.

Statistic	Mean	Median	St. Dev.
Volume (1000 $)	398.66	25	1,244.79
Trade costs (bps.)	10.82	0.00	37.41
Credit rating	9.53	9	4.42

the IRC trade cost measure, expressed as a spread in basis points over the mid-point price, on a relationship measure. Let Spread_{ijkbt} be the spread on a transaction that takes place at date-time t in bond b where dealer i sells to dealer j who then sells the security to dealer k within 15 minutes. We refer to the selling dealer as the seller, the intermediating dealer (dealer j) as the intermediating dealer (or dealer for short), and the buying dealer as the buyer. We measure the relationship strength between i, k, and j as

$$\text{Relationship Strength}_{ijkt} = \frac{1}{2}\frac{\text{Rolling volume}_{ij,t-31,t-1} + \text{Rolling volume}_{jk,t-31,t-1}}{\text{Rolling inter-dealer volume}_{j,t-31,t-1}} \tag{3.5.2}$$

Where in equation 3.5.2 $\text{Rolling volume}_{ij,t-31,t-1}$ is the rolling trade volume between dealers i and j from day $t-31$ to day $t-1$. Effectively we lag the proportion of trades by one day as it is known in the literature that trade size can impact trade costs.

We regress the inter-dealer IRC spread on this relationship measure as well as controls and fixed effects. We include a volume control due to the well known relationship between trade size and trade costs in the US corporate bond market. We include the log of the total volume of the contra parties to control for contra-party sophistication. We expect that more sophisticated or larger contra-parties will have lower trade costs. Our fixed effects include time (at the month, week or day level), bond, and dealer. We also include interacted fixed effects.

In our strictest specification we include triple interacted day times bond times intermediating dealer. In these specifications we are only looking at differences in spreads on trades in the same bond traded in the same day by the same intermediating dealer. We use two-way clustering of the standard errors at the bond and month level. This allows for correlation between any observation within the same month and for correlation in spreads in the same bond across months.

The results of these regressions are in table 3.5.2. We find that a one standard deviation increase in relationship strength is associated with 3 to 13 bps reduction in trade costs. The mean inter-dealer trade costs is approximately 11 bps. So, this represents between 0.2727 and 1.18 standard deviation decrease in trade costs. In tables 3.5.3 and 3.5.4 we perform the regressions of table 3.5.2 using the alternative relationship measures. In all specifications we find that stronger relationships predict reduced trade costs, that's significant at the 1% level.

Identifying the benefits of relationships is challenging as the decision to trade with a counterparty is a function of the terms of trade that the counterparty offers. If a counterparty experiences an idiosyncratic increase in trading costs, they will offer relatively poor terms for the dealer, making the dealer less likely to trade with them. This would imply that a dealer is likely to trade with counterparties whom

3.5.2. This table contains the results of the following regression Trade $\text{costs}_{ijkbt} = \beta\text{Relationship Strength}_{ijkt} + \text{Controls} + \text{F.E.} + \varepsilon_{ijkbt}$. Controls include the log of the par volume of the trade and the log of the average total volume that contra parties intermediate. The relationship measure are defined in the text. We exclude all trades by non-FINRA members and the relationship measures are only defined after a dealer has been in the sample for more than 30 days. Our sample runs from June 2002 until December 2016. The numbers number of fixed effects is in parenthesis in the fixed effect column. Standard errors are clustered at the bond and month level.

Dependent Variable:	Trade costs (bps.)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
(Intercept)	38.38***				
	(3.529)				
Proportion periods trade	-14.92***	-15.44***	0.1184	-18.38***	-1.425***
	(1.05)	(0.6954)	(0.7390)	(0.6841)	(0.1256)
log(Volume)	-0.0278	-1.024***	-4.343***	-1.474***	-0.6362***
	(0.0918)	(0.1278)	(0.3217)	(0.1941)	(0.0613)
log(Mean contra volume)	-0.8249***	0.3175***	-0.9034***	0.1279**	-0.0858***
	(0.1401)	(0.0512)	(0.0898)	(0.0618)	(0.0181)
Fixed-effects					
Bond (42,085)	No	Yes	No	No	No
Dealer (1,052)	No	No	Yes	No	No
Time (daily) (3,635)	No	No	No	Yes	No
Bond×Dealer×Time (daily) (6,093,638)	No	No	No	No	Yes
Fit statistics					
Observations	7,866,974	7,866,974	7,866,974	7,866,974	7,866,974
R^2	0.01556	0.19464	0.31105	0.09913	0.98407
Within R^2	–	0.0177	0.02855	0.02559	0.0035

Two-way (Bond & Time (monthly)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

they don't have a relationship with when the counterparties with whom they have a relationship with are offering poor terms of trade. This would imply that the estimates in our regressions underestimate the value of relationships to dealers.

3.5.3. This table contains the results of the following regression Trade $\text{costs}_{ijkbt} = \beta\text{Relationship Strength}_{ijkt} + \text{Controls} + \text{F.E.} + \varepsilon_{ijkbt}$. Controls include the log of the par volume of the trade and the log of the average total volume that contra parties intermediate. The relationship measure the average of whether the intermediaries traded in the previous month. We exclude all trades by non-FINRA members and the relationship measures are only defined after a dealer has been in the sample for more than 30 days. Our sample runs from June 2002 until December 2016. The numbers number of fixed effects is in parenthesis in the fixed effect column. Standard errors are clustered at the bond and month level.

Dependent Variable:	Trade costs (bps.)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
(Intercept)	76.8***				
	(4.681)				
Trade_last_month	-46.35***	-34.16***	-3.352***	-40.8***	-3.53***
	(1.441)	(0.8898)	(0.6378)	(1.082)	(0.6103)
log(Volume)	-0.0741	-0.9690***	-4.346***	-1.344***	-0.6195***
	(0.0943)	(0.1421)	(0.3239)	(0.2007)	(0.0610)
log(Mean contra volume)	-0.8979***	0.0816	-0.8846***	-0.1268*	-0.1027***
	(0.1241)	(0.0519)	(0.0819)	(0.0670)	(0.0186)
Fixed-effects					
Bond (42,089)	No	Yes	No	No	No
Dealer (1,059)	No	No	Yes	No	No
Time (daily) (3,635)	No	No	No	Yes	No
Bond×Dealer×Time (daily) (6,093,710)	No	No	No	No	Yes
Fit statistics					
Observations	7,867,693	7,867,693	7,867,693	7,867,693	7,867,693
R^2	0.01339	0.18741	0.31126	0.0888	0.98406
Within R^2	–	0.00883	0.02861	0.01412	0.00328

Two-way (Bond & Time (monthly)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

3.6 The Model

Motivated by the stylized facts we propose a model of inter-dealer trades. The model is set in discrete time, which is labeled with subscript $t \in \{0, 1, 2, \ldots\}$, and runs forever. There is a single asset, which we may refer to as a bond, that is in

3.5.4. This table contains the results of the following regression Trade $\text{costs}_{ijkbt} = \beta\text{Relationship Strength}_{ijkt} + \text{Controls} + \text{F.E.} + \varepsilon_{ijkbt}$. Controls include the log of the par volume of the trade and the log of the average total volume that contra parties intermediate. The relationship measure the average of whether the intermediaries traded in the previous month. We exclude all trades by non-FINRA members and the relationship measures are only defined after a dealer has been in the sample for more than 30 days. Our sample runs from June 2002 until December 2016. The numbers number of fixed effects is in parenthesis in the fixed effect column. Standard errors are clustered at the bond and month level.

Dependent Variable:	Trade costs (bps.)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
(Intercept)	34.86***				
	(3.685)				
Number of continuous trading months	-0.1515***	-0.0926***	-0.0559***	-0.0905***	-0.0144***
	(0.0147)	(0.0074)	(0.0064)	(0.0080)	(0.0013)
log(Volume)	-0.5302***	-1.159***	-4.401***	-1.405***	-0.6266***
	(0.1451)	(0.1611)	(0.3265)	(0.2042)	(0.0608)
log(Mean contra volume)	-0.5188***	0.1402***	-0.6685***	-0.1561**	-0.0835***
	(0.0845)	(0.0539)	(0.0682)	(0.0635)	(0.0186)
Fixed-effects					
Bond (42,102)	No	Yes	No	No	No
Dealer (1,070)	No	No	Yes	No	No
Time (daily) (3,635)	No	No	No	Yes	No
Bond×Dealer×Time (daily) (6,109,964)	No	No	No	No	Yes
Fit statistics					
Observations	7,889,970	7,889,970	7,889,970	7,889,970	7,889,970
R^2	0.02309	0.18751	0.31291	0.08558	0.98406
Within R^2	–	0.009	0.03072	0.01065	0.00339

Two-way (Bond & Time (monthly)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

positive net supply. The model features both customers and dealers. Customers are endowed with $z \in \{0, 1\}$ units of the asset when they arrive at the market and have a valuation δ_z, where z is their endowment and $\delta_0 > \delta_1 > 0$. Each period $N \geq 1$ new customers arrive at the market and are matched with one and only one dealer whom they may trade a quantity c with. The customer's life-time utility is given by

$$U = \delta_i(c+z) - P_c(c)c. \tag{3.6.1}$$

Where c is the quantity the customer and dealer trade, z is the initial endowment of the asset, and $P(c)$ is the dealer-to-customer price. The customers position, after trading with the dealer, is restricted to be in the *interval* $[0,1]$, i.e. $c+z \in [0,1]$.

In addition to customers our model features $2N$ dealers. Unlike customers dealers are infinitely lived and discount the future at rate $\frac{1}{1+r}$. All dealers have the same valuation for the asset, δ_d, $\delta_0 > \delta_d > \delta_1$. Dealers match with customers at random and this probability is independent across dealers and time. When a dealer matches with a customer they bargain over the terms of trade. The price the dealer-customer price is determined by Nash bargaining, where $\theta \in [0,1]$ is the dealers bargaining weight.

To intermediate a non-zero quantity c with a customer a dealer must pay a cost

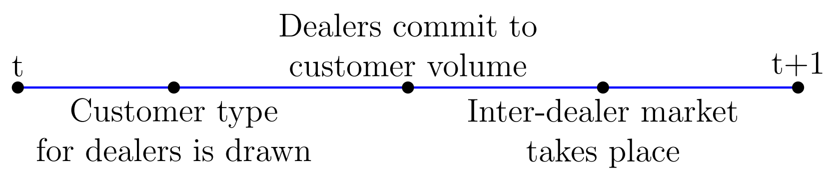
$$\alpha(c-X)^2 + f^c, \tag{3.6.2}$$

where $\alpha > \max_i |\delta_i - \delta_d|$, $f^c \geq 0$, $f^c < \min_i \frac{(\delta_i-\delta_d)^2(3-2\phi)}{2\alpha(2-\phi)^2}$, and X is any quantity the dealer intermediates on the inter-dealer market. $c - X$ is the final position of the dealer. We think of f^c as the fixed cost of matching with a customer. This includes the back office costs associated with transactions and FINRA fees. If the dealer and customer fails to come to an agreement the dealer receives no surplus for the period.

Within the same period a dealer may initiate a trade with another dealer. They may only trade with dealers who did not randomly match with another customer.[4] There are no search frictions in the inter-dealer market. We assume of simplicity that matching between a dealer with a customer and a dealer without a customer is random [5]. Prices in the inter-dealer market are determined by Nash bargaining. Where ϕ is the market power of the dealer who found the customer. We assume that for a dealer without a customer their are no fixed costs associated with the transaction. Thus, the cost for a dealer without a customer to intermediate a volume X is

$$\alpha(X)^2. \tag{3.6.3}$$

In figure 3.6.1 shows the timing of the model.



3.6.1. This figure presents the order of markets within a period of our model.

3.6.1 Model Results

Let $P^d(X)$ be the inter-dealer price. We solve our problem backwards. Consider a dealer who has found a customer and has committed to a transaction of size C with a customer. We will solve for the case where the dealer is selling the asset to the customer. The case where it is a customer purchase is similar. When the dealer

[4] This is to simplify the problem of a dealer who matches with a customer, as they receive no utility for the period if they match with a customer but fail to trade with them.

[5] This is somewhat unrealistic given we have shown that the probability that a dealer pair transacts depends on the history of transactions between the pair. This is perhaps not so problematic when N is small. We leave improving on this matching technology for future research.

with the customer enters the inter-dealer market they can purchase from the dealer. The payoff for the dealer is thus

$$\pi(c, X) = \begin{cases} cP_i(c) - XP^d(X) - \delta_d(c - X) - f^c - \alpha(c - X)^2, c \neq 0 \\ 0, c = 0. \end{cases} \tag{3.6.4}$$

For the dealer without a customer their payoff is given by

$$\pi(X) = XP^d(X) - \delta_d X - \alpha(X)^2. \tag{3.6.5}$$

Since we are looking a customer sale we would have $X > 0$. For the case where it is a customer buy we would have $X < 0$. Since the marginal cost functions are identical the optimal transaction volume is $c/2$ and the maximized joint surplus of the dealers is $\alpha\frac{c^2}{2}$.[6] Given that the dealer who found the customer receives a fraction ϕ of the surplus the inter-dealer price at the optimal quantity X^* is

$$P^d(X^*) = \delta_d + \alpha\left(\frac{3}{2} - \phi\right)c.^7 \tag{3.6.6}$$

When the dealer with a customer purchases the asset from the first dealer in a sense they receive a discount. The size of this discount depends on the size of ϕ, their share of the surplus.

We now move backwards to the first stage of the period, that is where the dealers are matched with customers. Using the previous results it is easy to see that for a

---

[6]See lemma 1 in appendix 3.C for details

[7]See lemma 2

dealer who has a customer their payoff is given by

$$\Pi(c) = (P_i(c) - \delta_d)\, c - f_c - \alpha(c)^2 + \phi\alpha(c)^2/2. \tag{3.6.7}$$

The dealer and customer maximize the joint surplus. Thus the optimal quantity for the pair to intermediate is

$$c^* = \frac{\delta_i - \delta_d}{\alpha(2-\phi)},^{8} \tag{3.6.8}$$

where δ_i is the valuation of the customer ($i = 0$ in our case as the customer is a buyer). We can also write the dealer-customer price as

$$P_i(c) = \theta\delta_i + (1-\theta)\left[\delta_d + \frac{f^c}{c} + \alpha c - \phi\alpha\frac{c}{2}\right]. \tag{3.6.9}$$

Using 3.6.9 we can show that the dealer-customer price at the equilibrium quantity equals,

$$P_i^* = \theta\delta_i + (1-\theta)\left[\frac{1}{2}(\delta_d + \delta_i) + \frac{f^c\alpha(2-\phi)}{\delta_i - \delta_d}\right].^{9} \tag{3.6.10}$$

3.6.2 Model Discussion

The socially optimal quantity for dealers and customers to intermediate is

$$c^{Opt} = \frac{\delta_i - \delta_d}{\alpha},^{10} \tag{3.6.11}$$

[8] See lemma 3

[9] See lemma 4

[10] See lemma 5

which is strictly less than the equilibrium amount in 3.6.8 for $\phi < 1$. This is because of double marginalization. In the inter-dealer stage of the period the pair of dealers maximize the joint surplus. However, the dealer who finds the customer only receives a fraction ϕ of the surplus. The dealer and customer take into account the fact that they only receive a fraction of the surplus from the inter-dealer market. In anticipation of receiving only a small fraction of this surplus the dealer and customer choose a smaller quantity when they meet, leading to inefficiency. The loss in total surplus equals

$$\text{Lost surplus} = \frac{(\delta_i - \delta_d)^2}{2\alpha}\left(\frac{1-\phi}{2-\phi}\right)^2, ^{11} \tag{3.6.12}$$

which is positive for $\phi < 1$, reflecting the losses from the double marginalization externality.

Our model predicts that there is a u-shaped relationship (for $f^c > 0$) between customer transaction size and transaction cost. For small transactions prices fall as transaction volume rises, as we see in the data. However, at larger volume our model predicts that as transaction volume rises the customer transaction costs will also rise. Our regression, as well as most research suggests that customer enjoy a size discount, larger transactions being less costly.[12].

Our model features random matching between the dealers in the inter-dealer market. This means that the probability of trading between the dealers is constant and the same across dealers. While this is empirically counterfactual it greatly simplifies our analysis.

[11]See lemma 6

[12]See Edwards et al. (2004) for example. However, Pintér et al. (2022) finds that once one controls for the customers identity their is a size premium in the UK corporate bond market.

3.6.3 Relationship Among Dealers

Since we have random matching in the inter-dealer market we will examine non-exclusive relationships (for $N > 1$) or coalitions involving all dealers. When in a relationship all dealers who find customers commit to intermediating the socially optimal amount with customers. When a dealer matches with a customer they receive a total of $\theta\left(\frac{(\delta_i-\delta_d)^2}{2\alpha(2-\phi)} - f^c\right)$ in surplus. If they choose to intermediate the optimal amount they receive $\theta\left(\frac{\phi(\delta_i-\delta_d)^2}{2\alpha} - f^c\right)$.[13] Hence, their surplus is lower for $\phi < 1$. For a dealer who doesn't have a customer their surplus at quantity c^* is $(1-\phi)\left(\frac{\delta_i-\delta_d}{2(2-\phi)}\right)$ which is less than $(1-\phi)\left(\frac{\delta_i-\delta_d}{2}\right)$, which is larger for $\phi < 1$. Summing across both dealer's their surplus is larger by amount $\theta\frac{(\delta_i-\delta_d)^2}{2\alpha}\left(\frac{1-\phi}{2-\phi}\right)^2$. Thus by a version of the folk theorem for infinitely repeated games there exists a discount rate $\frac{1}{1+r}$ that sufficiently small where we can support our relationship equilibrium.[14]

Our model makes a number of predictions regarding the impact of relationships on prices and quantities traded. Comparing the equilibrium without relationships to one with we can see that the quantity traded between the dealers is larger when the dealers are in a relationship. We also note that the average customer transaction volume is larger in a market with strong relationships compared to one without relationships.

Our model raises several issues in terms of measuring relationship strength. The random matching assumption of our model means that the pattern of past trades does not predict a relationship between dealers. This of course makes it difficult to identify the existence of relationships from data. We leave refining this assumption

[13]Relationships can also be modelled as the dealers agreeing to let $\phi = 1$.

[14]See Fudenberg and Tirole (1991)

for future research.

3.7 Trading Volume

3.7.1. This table contains the customer trading volume summary statistics. We exclude the first 90 days of trading to avoid the on-the-run phenomenon. We include only complete months.

Statistic	Mean	Median	St. Dev.
Mean trade size	860,840.20	213,655.10	1,954,137.00
Number of trades	34.18	12	86.96
log(Volume)	14.56	14.97	2.79
Relationship measure	0.09	0.02	0.16

Our model suggests that relationships should also increase the total customer trading volume. Bond markets where the dealer have stronger relationships should see increased trading volume from customers. Empirically, there are two potential channels for this. The first is that each customer increases the volume they wish to trade. This would imply that the average trade size should increase for customer who would previously trade in the market. The second mechanism, which is not present in our model, is that when trading costs fall more customer's with lower valuations enter the market and begin trading. This would lead to an increase in the total number of customer trades.

An issue with examining how relationships effect the mean customer trade size is that the marginal customers who participate in the may have a smaller trading needs. If relationships increased customer participation in the market then if the marginal customer has a smaller trading needs than the mean it would tend to decrease the mean trade size, even if the trade size for existing customer increased. Letting the Vol_{bcm} be the total volume of customer c in month m we define this for the customers who participate in the market as well as those that choose not to participate. For customers who choose not to participate in the market we have that $Vol_{bcm} = 0$.

Define $T_{cmb}(Vol)$ as a function that maps customer trading volume to total trading costs and $U_c(Vol)$ as the customer's utility function that maps the trading volume to their utility for the month. Assume customers exist for one period and that trading volume in a given month is separable from the other consumption goods of the customer. As is true in the data assume $T_{cmb}(.)$ is increasing, continuous, and concave and $U_c(.)$ is increasing, continuous, and concave as well. The optimality conditions imply that the optimal vol^*_{cmb} solves

$$vol^*_{cmb} = \begin{cases} U_c'^{-1}(T'_{cmb}(vol^*_{cmb})) \quad \text{if} \quad U_c(0) < U_c(vol^*_{cmb}) \\ 0 \text{ otherwise} \end{cases} \tag{3.7.1}$$

As relationships improve we expect that the trading cost schedule offered to customers to improve. We mean this in the sense that the offer curve $T_{cmb}(.)$ will shift down. This will cause an increase in the volume traded of customers who already have a positive volume but will also increase participation of customers who's volume is currently zero. Since we don't observe the customers who are currently not trading and the customer identities are not available in our data the effect on the measured mean customer volume is ambiguous. In general it will depend on the distribution of customer utility functions for those that participate in the market and those that don't, which is of course unobservable.

To empirically test these theories we aggregate the data up to the monthly level.

We calculate the number of customer trades, defined as the total number of customer trades in the bond that month. The total customer trading volume, defined as the sum of the total par volume of all customer trades. The customer turnover defined as the total customer trading volume divided by the total amount outstanding of the bond. We defined the mean trade size as the mean par volume across all customer trades in the bond that month.

We regress these variables on measures of relationship strength at the bond level. One measure we propose is the proportion of trades completed between two dealers in the previous 30 days and then aggregating this to the bond/month level by taking the weighted mean of this measure weighting by par volume traded in bond b for each dealer. We use this measure of relationship strength to perform the following regressions

$$y_{bm} = \beta \text{Relationship}_{bm} + \text{F.E.} + \varepsilon_{bm}. \tag{3.7.2}$$

Where the y_{bm} on the left-hand-side of 3.7.2 represents the customer trading volume measures discussed above. Our fixed effects include bond and monthly fixed effects. For these regressions we exclude the first 90 days after the bond is issued and only include complete months. This is because trading volume is significantly higher when a bond is newly issued and incomplete months, due to maturity would lead tend to artificially depress trading volume. Summary statistics are in table 3.7.1

In table we have 3.7.2 we regress the monthly number of customer trades in a given bond on the relationship measure. We find that a one standard deviation increase in the bond specific relationship measure leads to an increase in the monthly number of customer trades of between 3.2 and 7.7.

In our next set of results in table 3.7.3 we regress the log of the total customer volume on our relationship measures. As our model predicts an increase in the relationships increase the total customer trading volume. In our specification with bond and monthly fixed effects we find that a one standard deviation increase in the relationship measure leads to a 5.6% increase in customer volume in the market.

3.7.2. This table contains the results of the following regression $y_{bm} = \beta \text{Relationship}_{bm} + \text{F.E.} + \varepsilon_{bm}$ where y_{bm} is the total number of customer trades in a bond that month. We exclude the first 90 days of trading to avoid the on-the-run phenomenon. We include only complete months. All standard errors are double clustered at the bond and month level.

Dependent Variable:	Number of trades per month			
Model:	(1)	(2)	(3)	(4)
Variables				
(Intercept)	36.3***			
	(0.67)			
Bond relationship measure	48***	47.9***	20.9***	20.1***
	(1.9)	(1.9)	(0.81)	(0.75)
Fixed-effects				
Time (monthly)		Yes		Yes
Bond			Yes	Yes
Fit statistics				
Observations	1,662,081	1,662,081	1,662,081	1,662,081
R^2	0.00602	0.01347	0.49885	0.50619
Within R^2	–	0.00602	0.00189	0.00177

Two-way (Bond & Time (monthly)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

3.7.3. This table contains the results of the following regression $y_{bm} = \beta\text{Relationship}_{bm} + \text{F.E.} + \varepsilon_{bm}$ where y_{bm} is the log of customer volume in a bond that month. We exclude the first 90 days of trading to avoid the on-the-run phenomenon. We include only complete months. All standard errors are double clustered at the bond and month level.

Dependent Variable:	log(Customer volume)			
Model:	(1)	(2)	(3)	(4)
Variables				
(Intercept)	14.7***			
	(0.04)			
Bond relationship measure	1.2***	1.3***	0.33***	0.35***
	(0.06)	(0.06)	(0.02)	(0.02)
Fixed-effects				
Time (monthly)		Yes		Yes
Bond			Yes	Yes
Fit statistics				
Observations	1,662,081	1,662,081	1,662,081	1,662,081
R^2	0.00493	0.02974	0.76541	0.78446
Within R^2	–	0.00561	0.00126	0.00151

Two-way (Bond & Time (monthly)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

3.8 Conclusion

In this paper we show that trading patterns in the US corporate bond market resemble trading relationships rather than a series of spot markets. Using the academic enhanced version of the TRACE data we demonstrate empirically that the dealers only trade with a small number of other dealers rather than with the market as a whole. We also show that the group of dealers whom they trade with is highly persistent over time. Motivated by this we propose several measures for relationship strength, including the proportion of inter-dealer trading volume between the pair,

3.7.4. This table contains the results of the following regression $y_{bm} = \beta\text{Relationship}_{bm} + \text{F.E.} + \varepsilon_{bm}$ where y_{bm} is the total customer volume divided by the amount outstanding for the bond that month. We exclude the first 90 days of trading to avoid the on-the-run phenomenon. We include only complete months. All standard errors are double clustered at the bond and month level.

Dependent Variable:	Turnover			
Model:	(1)	(2)	(3)	(4)
Variables				
(Intercept)	0.04***			
	(0.0007)			
Bond relationship measure	0.04***	0.04***	0.02***	0.02***
	(0.002)	(0.002)	(0.001)	(0.001)
Fixed-effects				
Time (monthly)	No	Yes	No	Yes
Bond	No	No	Yes	Yes
Fit statistics				
Observations	1,570,588	1,570,588	1,570,588	1,570,588
R^2	0.00339	0.01405	0.35162	0.36323
Within R^2	–	0.00378	0.00112	0.00123

Two-way (Bond & Time (monthly)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

the number of continuous months the pair have been trading, and whether the pair have traded in the previous 30 days. We show that these measures are positively correlated with each other.

We measure dealer transaction costs using the implied round trip cost for inter-dealer transactions. Using the relationship measures we developed we show that inter-dealer transaction costs are lower when the pair of dealers have stronger relationships. In our strictest specification we include date (monthly) interacted with bond interacted with dealer fixed effects to control for time varying differences in

3.7.5. This table contains the results of the following regression $y_{bm} = \beta \text{Relationship}_{bm} + \text{F.E.} + \varepsilon_{bm}$ where y_{bm} is the average customer trade size in the bond market. We exclude the first 90 days of trading to avoid the on-the-run phenomenon. We include only complete months. All standard errors are double clustered at the bond and month level.

Dependent Variable:	Mean par volume of trades			
Model:	(1)	(2)	(3)	(4)
Variables				
(Intercept)	798175***			
	(7819.7)			
Bond relationship measure	-153218***	-140448.4***	-97819.3***	-81814.1***
	(21135.7)	(21222.9)	(9498.1)	(10084.6)
Fixed-effects				
Time (monthly)		Yes		Yes
Bond			Yes	Yes
Fit statistics				
Observations	1,662,081	1,662,081	1,662,081	1,662,081
R^2	0.00024	0.00348	0.45759	0.47305
Within R^2	–	2e-04	0.00015	0.00011

Two-way (Bond & Time (monthly)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

bond transactions costs by dealer. We also show that when dealers have stronger relationships with each other their bilateral trading volume is larger.

Motivated by our empirical results we construct a model of trading relationships in the US corporate bond market. In our model some dealers first locate a customer and commit to transacting a particular volume with them. After locating a customer, they can trade with the dealers who don't locate customers. There are gains to trade because the cost of fulfilling customer's order is convex in volume. However, market power in the inter-dealer market leads to double marginalization. This leads to inefficiency as when the dealer finds a customer, they optimally decide to intermediate a smaller volume as they only receive a part of the surplus when they move to inter-dealer market.

In our model dealers repeatedly trade with each other. This allows dealers to develop a reputation or relationships. When dealers have relationships the dealer who locates a customer will commit to trading the optimal amount with a customer. Total surplus for dealers is larger and since the probability of dealer locating a customer is constant across periods. Thus, the existence of relationships restores optimality. Using our model, we show that average customer trading volume is larger when dealers have strong relationships with each other. We validate this prediction showing that customer trading volume is larger when dealers have stronger relationships with each other.

APPENDIX

3.A Price Error Filter

In this section we describe the price error filer that we implement in cleaning our data. We follow Rossi (2014). For each trade we calculate a left and right rolling median price using the nearest 20 trades on each side provided that they occur within 20 days. We then calculate the absolute deviation of the reported price from these rolling medians. For each week we then calculate the 99% percentile of the rolling deviations amongst all trades. If both the left and right rolling deviations are larger than these cutoffs we consider flag the reported transaction as an outlier. For transactions where the prices were flagged as outliers we replace the reported price with the rolling median defined as the the midpoint between the left and right rolling medians.

3.B Constructing the IRC Trade Costs Measure

One method we use to measure trade costs follows Feldhütter (2012). In that paper the author defines round trip customer trading costs as a customer to dealer and then dealer to customer transaction that takes place in a short window of time on the same bond, with the same dealer and with the same volume. We employ a similar method as in this paper. However, as our focus is on inter-dealer trading costs, we define trading costs as the difference in the average sale and average buy price of the middle dealer in matched a dealer-to-dealer-to-dealer (DDD) transaction. A matched trade is trade where the dealer reverses the transaction within 15 minutes.

That is there is a series of trades on the same bond, in the same day, by the same dealer, where the time difference between the subsequent transactions is no longer than 15 minutes such that change in the dealer's cumulative inventory is zero.

Usually this average is just a single buy/sell price, but in some cases there can be multiple transactions within the same window. Here, we match the transaction based on the bond, the volume traded, the identity of the middle dealer. We also impose the difference in the time between the buy and the sell must not be more than 15 minutes, reflecting the TRACE 15 minute reporting window that bond dealers have.

After matching these trades we calculate the cumulative change in the signed inventory of the middle dealer. At the first point during that 15 minute window the change in inventory is zero we group all trades together into a single group. We calculate the average selling and buying price across these transactions. The difference between these two is the average trade cost We then start the group again. If the signed inventory volume reaches zero again then all the transactions in this interval are another group, etc. If during the 15 minute window the signed volume never reaches zero, then we exclude all these transactions in this group.

3.C Proofs

Lemma 1. *The optimal quantity to intermediate in the inter-dealer market is* $\frac{c}{2}$ *and the joint surplus is* $\alpha\frac{c^2}{2}$

Proof. Following section 3.6.1 Assume we are looking at a customer buy case. The total surplus for the dealers is $cP(c) - \delta_d c - f^c - \alpha(c^2 - 2cX + 2X^2)$. Differentiation reveals that $X^* = \frac{c}{2}$. The surplus from the inter-dealer transaction for quantity X^* is $cP(c) - \delta_d c - f^c - \alpha(\frac{c^2}{2})$ subtracting $cP(c) - \delta_d c - f^c - \alpha(c^2)$ yields $\frac{c^2}{2}$

□

Lemma 2. *The inter-dealer price at the optimal quantity, $X^* = \frac{c}{2}$, is $P^d(X^*) = \delta_d + \alpha\left(\frac{3}{2} - \phi\right)c$*

Proof. The dealer who found the customer receives a fraction ϕ of the total surplus from the inter-dealer trade. Thus, the difference in surplus from finding entering the inter-dealer market and not entering the secondary market must equal $\phi\alpha\frac{c^2}{2}$. This difference in surplus is $-XP^d(X) + \delta_d X - \alpha(X^2 - 2cX)$. Plugging in $X^* = \frac{c}{2}$ and equating to $\phi\alpha\frac{c^2}{2}$ yields the stated result after simplification.

□

Lemma 3. *The optimal quantity to intermediate when a dealer and customer meet is $\frac{\delta_i - \delta_d}{\alpha(2-\phi)}$ and the joint surplus of the dealer and customer is $\frac{(\delta_i - \delta_d)^2}{2\alpha(2-\phi)} - f^c$*

Proof. Following section 3.6.1 Assume we are looking at a customer buy case. The total surplus for the dealer and customer is $(\delta_i - \delta_d)c - f^c - \alpha c^2 + \phi\alpha\frac{c^2}{2}$. Differentiation reveals that $c^* = \frac{\delta_i - \delta_d}{\alpha(2-\phi)}$. Inserting c^* into the total surplus function reveals $\frac{(\delta_i - \delta_d)^2}{2\alpha(2-\phi)} - f^c$.

□

Lemma 4. *The dealer-customer price is $P_i(c) = \theta\delta_i + (1-\theta)\left[\delta_d + \frac{f^c}{c} + \alpha c - \phi\alpha\frac{c}{2}\right]$ and at quantity $c^* = \frac{\delta_i - \delta_d}{\alpha(2-\phi)}$ it equals $P_i^* = \theta\delta_i + (1-\theta)\left[\frac{1}{2}(\delta_d + \delta_i) + \frac{f^c\alpha(2-\phi)}{\delta_i - \delta_d}\right]$*

Proof. The dealer receives a fraction θ of the surplus. If they do not accept a customer they receive no surplus. Thus, from equation 3.6.7, their surplus equals

$(P_i(c) - \delta_d)\, c - f_c - \alpha(c)^2 + \phi\alpha(c)^2/2$. Setting this equal to their share of the total surplus, $\theta\left[(\delta_i - \delta_d)c - f^c - \frac{\alpha}{2}c^2(2-\phi)\right]$ and rearranging yields $P_i(c) = \theta\delta_i + (1-\theta)\left[\delta_d + \frac{f^c}{c} + \alpha c - \phi\alpha\frac{c}{2}\right]$. Inserting c^* and simplifying completes the proof.

□

Lemma 5. *The socially optimal quantity is* $c^{Opt} = \frac{\delta_i - \delta_d}{\alpha}$.

Proof. From lemma 1 we know that the total surplus for all parties is $(\delta_i - \delta_d)c - f^c - \frac{\alpha}{2}c^2$. Differentiation this equation and setting it equal to zero yields the result. □

Lemma 6. *The loss in total surplus arising from double marginalization is*

$$\frac{(\delta_i-\delta_d)^2}{2\alpha}\left(\frac{1-\phi}{2-\phi}\right)^2.$$

Proof. Again, from lemma 1, we know that the total surplus for all parties is $(\delta_i - \delta_d)c - f^c - \frac{\alpha}{2}c^2$. The difference between this at c^{Opt} and c^* is $(\delta_i - \delta_d)(c^{Opt} - c^*) - \frac{\alpha}{2}((c^{Opt})^2 - (c^*)^2)$. We have that $c^{Opt} - c^* = \frac{(\delta_i-\delta_d)}{\alpha}\left(1 - \frac{1}{2-\phi}\right) = \frac{(\delta_i-\delta_d)}{\alpha}\left(\frac{1-\phi}{2-\phi}\right)$ and $(c^{Opt})^2 - (c^*)^2 = \frac{(\delta_i-\delta_d)^2}{\alpha^2}\left(1 - \frac{1}{(2-\phi)^2}\right) = \frac{(\delta_i-\delta_d)^2}{\alpha^2}\left(\frac{(3-\phi)(1-\phi)}{(2-\phi)^2}\right)$. Inserting these into our equation and simplifying yields $\frac{(\delta_i-\delta_d)^2}{2\alpha}\left(\frac{1-\phi}{2-\phi}\right)^2$

□

Bibliography

Afonso, Gara, Anna Kovner, and Antoinette Schoar, 2014, Trading partner in the interbank market, Federal Reserve Bank of New York Staff Report no. 620.

Aguirregabiria, Victor, and Pedro Mira, 2002, Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models, *Econometrica* 70, 1519–1543.

Ahlfeldt, Gabriel M., Fabian Bald, Duncan Roth, and Tobias Seidel, 2020, The stationary spatial equilibrium with migration costs, Technical report.

Almagro, Milena, and Tomás Domınguez-Iino, 2021, Location sorting and endogenous amenities: Evidence from amsterdam, *NYU, mimeograph* .

An, Yu, 2019, Competing with inventory in dealership markets, Working paper, Johns Hopkins.

Anderson, Simon P., Andre De Palma, and Jacques-Francois Thisse, 1992, *Discrete choice theory of product differentiation* (MIT press).

Arcidiacono, Peter, and Paul B. Ellickson, 2011, Practical methods for estimation of dynamic discrete choice models, *Annual Review of Economics* 3, 363–394.

Arcidiacono, Peter, and Robert A. Miller, 2011, Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity, *Econometrica* 79, 1823–1867.

Atkeson, Andrew, Andrea Eisfeldt, and Pierre-Olivier Weill, 2015, Entry and exit in otc derivatives markets, *Econometrica* 83, 2231–2292.

Babus, Ana, and Tai-Wei Hu, 2018, Endogenous intermediation in over-the-counter markets, *Journal of Financial Economics* 125, 200–215.

Babus, Ana, and Peter Kondor, 2018, Trading and information diffusion in otc markets, *Econometrica* 86, 1727–1769.

Baer, Justin, 2020, The Day Coronavirus Nearly Broke the Financial Markets, *The Wall Street Journal* , May 20.

Bao, Jack, Maureen O'Hara, and Alex Zhou, 2018, The Volcker rule and market making in times of stress, *Journal of Financial Economics* 130, 95–113.

Barkai, Simcha, 2020, Declining labor and capital shares, *The Journal of Finance* 75, 2421–2463.

Barrero, Jose Maria, Nicholas Bloom, and Steven J Davis, 2021, Why working from home will stick, Technical report, National Bureau of Economic Research.

Baum-Snow, Nathaniel, Matthew Freedman, and Ronni Pavan, 2018, Why has urban inequality increased?, *American Economic Journal: Applied Economics* 10, 1–42.

Bernanke, Ben, and Janet Yellen, 2020, The Federal Reserve must reduce long-term damage from coronavirus, *Financial Times* , March 18.

Berry, Steven T, 1994, Estimating discrete-choice models of product differentiation, *The RAND Journal of Economics* 242–262.

Bessembinder, Hendrik, Stacey Jacobsen, William Maxwell, and Kumar Venkataraman, 2018, Capital commitment and illiquidity in corporate bonds, *The Journal of Finance* 73, 1615–1661.

Bilal, Adrien, and Esteban Rossi-Hansberg, 2021, Location as an asset, *Econometrica* 89, 2459–2495.

Bowles, Nellie, 2021, They can't leave the bay area fast enough, *The New York Times* .

Boyarchenko, Nina, Anna Kovner, and Or Shachar, 2020, It's what you say and what you buy: A holistic evaluation of the corporate credit facilities, Working paper, FRB New York.

Brynjolfsson, Erik, John J Horton, Adam Ozimek, Daniel Rock, Garima Sharma, and Hong-Yi TuYe, 2020, Covid-19 and remote work: an early look at us data .

Bunten, Devin Michelle, 2017, Is the rent too high? aggregate implications of local land-use regulation .

Campbell, John Y, Stefano Giglio, and Parag Pathak, 2011, Forced sales and house prices, *American Economic Review* 101, 2108–31.

Chen, Jiakai, Haoyang Liu, Asani Sarkar, and Zhaogang Song, 2020, Cash-forward arbitrage and dealer capital in MBS makers: COVID-19 and beyond, Working paper, University of Hawaii, FRB, and Johns Hopkins University.

Choi, Jaewon, and Yesol Huh, 2018, Customer liquidity provision: Implications for corporate bond transaction costs, Working paper, UIUC and Federal Reserve Board.

Colliard, Jean-Edouard, Thierry Foucault, and Peter Hoffmann, 2018, Inventory management, dealers' connections, and prices in otc markets, Working paper.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, 2020, Income growth and the distributional effects of urban spatial sorting, *Working Paper* .

D'Amico, Stefania, Vamsidhar Kurakula, and Stephen Lee, 2020, Impacts of the fed corporate credit facilities through the lenses of ETFs and CDX, Working paper, FRB Chicago.

Darmouni, Olivier, and Kerry Y. Siani, 2020, Crowding out bank loans: Liquidity-driven bond issuance, Working paper, Columbia Business School.

Davis, Morris, and Jonathan Heathcote, 2007, The price and quantity of residential land in the united states, *Journal of Monetary Economics* 54, 2595–2620.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song, 2017, The value of trading relations in turbulent times, *Journal of Financial Economics* 124, 266–284.

Diamond, Rebecca, 2016, The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000, *American Economic Review* 106, 479–524.

Diamond, Rebecca, Tim McQuade, and Franklin Qian, 2019, The effects of rent control expansion on tenants, landlords, and inequality: Evidence from san francisco, *American Economic Review* 109, 3365–94.

Dick-Nielsen, Jens, 2014, How to clean enhanced TRACE data, Working paper, CBS.

Duffie, Darrell, 2020, Still the world's safe haven? Redesigning the US treasury market after the COVID-19 crisis, Hutchins Center Working paper #62.

Duffie, Darrell, Nicolae Gârleanu, and Lasse H. Pedersen, 2007, Valuation in over-the-counter markets, *The Review of Financial Studies* 20, 1865–1900.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen, 2005, Over-the-counter markets, *Econometrica* 73, 1815–1847.

Ebsim, Mahdi, Miguel Faria-e Castro, and Julian Kozlowski, 2020, Corporate bond spreads and the pandemic, Federal Reserve Bank of Saint Louis, On the Economy Blog, Posted April 9, 2020.

Edwards, Amy K., Lawrence E. Harris, and Michael S. Piwowar, 2004, Corporate bond market: Transparency and transaction costs, working paper.

Eisfeldt, Andrea, Bernard Herskovic, Srimam Rajan, and Emil N. Siriwardane, 2018, Otc intermediaries, Working paper, UCLA.

Fajgelbaum, Pablo D, and Cecile Gaubert, 2020, Optimal Spatial Policies, Geography, and Sorting, *The Quarterly Journal of Economics* 135, 959–1036.

Falato, Antonio, Itay Goldstein, and Ali Hortaçsu, 2020, Financial fragility in the COVID-19 crisis: The case of investment funds in corporate bond markets, Working paper, FRB, Wharton, and University of Chicago.

Farboodi, Maryam, 2014, Intermediation and voluntary exposure to counterparty risk, Working paper, MIT.

Favilukis, Jack, Pierre Mabille, and Stijn Van Nieuwerburgh, 2019, Affordable housing and city welfare .

Favilukis, Jack, and Stijn Van Nieuwerburgh, 2017, Out-of-town home buyers and city welfare, *The Journal of Finance* .

Feldhütter, Peter, 2012, The same bond at different prices: Identifying search frictions and selling pressures, *The Review of Financial Studies* 25, 1155–1206.

Flanagan, Thomas, and Amiyatosh Purnanandam, 2020, Corporate bond purchases after covid-19: Who did the fed buy and how did the markets respond?, Working paper, University of Michigan.

Fleming, Michael J., and Francisco Ruela, 2020, Treasury market liquidity during the COVID-19 crisis, Working paper, FRB New York.

Foley-Fisher, Nathan, Gary Gorton, and Stéphane Verani, 2020, The dynamics of adverse selection in privately-produced safe debt markets, Working paper, FRB and Yale School of Management.

Fudenberg, Drew, and Jean Tirole, 1991, *Game theory* (MIT press).

Ganong, Peter, and Daniel Shoag, 2017, Why has regional income convergence in the us declined?, *Journal of Urban Economics* 102, 76–90.

Gilchrist, Simon, Bin Wei, Vivian Z. Yue, and Egon Zakrajšek, 2020, The Fed Takes on Corporate Credit Risk: An Analysis of the Efficacy of the SMCCF, Working paper, NYU, FRB, Emory, and BIS.

Glaeser, Edward, and Joseph Gyourko, 2018, The economic implications of housing supply, *Journal of Economic Perspectives* 32, 3–30.

Glaeser, Edward L, Joseph Gyourko, and Raven E Saks, 2005, Why have housing prices gone up?, *American Economic Review* 95, 329–333.

Goldberg, Jonathan, and Yoshio Nozawa, 2020, Liquidity supply in the corporate bond market, *The Journal of Finance* , Forthcoming.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, 2020, Bartik instruments: What, when, why, and how, *American Economic Review* 110, 2586–2624.

Goldstein, Michael A., and Edith S. Hotchkiss, 2020a, Providing liquidity in an illiquid market: Dealer behavior in us corporate bonds, *Journal of Financial Economics* 135, 16–40.

Goldstein, Michael A, and Edith S Hotchkiss, 2020b, Providing liquidity in an illiquid market: Dealer behavior in us corporate bonds, *Journal of Financial Economics* 135, 16–40.

Gupta, Arpit, Vrinda Mittal, Jonas Peeters, and Stijn Van Nieuwerburgh, 2021, Flattening the curve: pandemic-induced revaluation of urban real estate, Technical report, National Bureau of Economic Research.

Gyourko, Joseph, Jonathan S Hartley, and Jacob Krimmel, 2021, The local residential land use regulatory environment across us housing markets: Evidence from a new wharton index, *Journal of Urban Economics* 124, 103337.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai, 2013, Superstar cities, *American Economic Journal: Economic Policy* 5, 167–99.

Gyourko, Joseph, and Raven Molloy, 2015, Regulation and housing supply, in *Handbook of regional and urban economics*, volume 5, 1289–1337 (Elsevier).

Gyourko, Joseph, Albert Saiz, and Anita Summers, 2008, A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index, *Urban Studies* 45, 693–729.

Haddad, Valentin, Alan Moreira, and Tyler Muir, 2020, When selling becomes viral: Disruptions in debt markets in the COVID-19 crisis and the Fed's response, *The Review of Financial Studies* , Forthcoming.

He, Zhiguo, Stefan Nagel, and Zhaogang Song, 2020, Treasury inconvenience yields during the covid-19 crisis, Working paper, University of Chicago and Johns Hopkins University.

Hendershott, Terrence, Dan Li, Dmitry Livdan, and Norman Schürhoff, Forthcoming, Relationship Trading in OTC Markets, *The Journal of Finance* .

Herkenhoff, Kyle, Lee Ohanian, and Edward Prescott, 2017, Tarnishing the golden and empire states: Land-use restrictions and the u.s. economic slowdown, *Journal of Monetary Economics* 93.

Hornbeck, Richard, and Enrico Moretti, 2019, Estimating who benefits from productivity growth: Direct and indirect effects of city manufacturing tfp growth on wages, rents, and inequality, Nber working papers, National Bureau of Economic Research.

Hotz, V Joseph, and Robert A Miller, 1993, Conditional choice probabilities and the estimation of dynamic models, *The Review of Economic Studies* 60, 497–529.

Hsieh, Chang-Tai, and Enrico Moretti, 2019, Housing constraints and spatial misallocation, *American Economic Journal: Macroeconomics* 11, 1–39.

Hugonnier, Julien, Ben Lester, and Pierre-Olivier Weill, 2019, Frictional intermediation in Over-the-Counter markets, *Review of Economic Studies* Forthcoming, Working paper EPFL, Federal Reserve Bank of Philadelphia, and UCLA.

Idzelis, Christine, 2020, The Corporate Bond Market Is 'Basically Broken,' Bank of America Says, *Institutional Investor* , March 19.

Jordá, Óscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor, 2019, The Rate of Return on Everything, 1870–2015*, *The Quarterly Journal of Economics* 134, 1225–1298.

Kennan, John, and James R. Walker, 2011, The effect of expected income on individual migration decisions, *Econometrica* 79, 211–251.

Kiyotaki, Nobuhiro, Alexander Michaelides, and Kalin Nikolov, 2011, Winners and losers in housing markets, *Journal of Money, Credit and Banking* 43, 255–296.

Kiyotaki, Nobuhiro, Alexander Michaelides, and Kalin Nikolov, 2020, Housing, distribution and welfare, *Working Paper* .

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, 2000, Capital-skill complementarity and inequality: A macroeconomic analysis, *Econometrica* 68, 1029–1053.

Krusell, Per, and Anthony A Smith, Jr, 1998, Income and wealth heterogeneity in the macroeconomy, *Journal of political Economy* 106, 867–896.

Kukura, Joe, 2018, Why a laundromat might be considered 'historic' .

Li, Dan, and Norman Schürhoff, 2019, Dealer networks, *The Journal of Finance* 74, 91–144, Working Paper, HEC Lausanne.

Ma, Yiming, Kairong Xiao, and Yao Zeng, 2020, Mutual fund liquidity transformation and reverse flight to liquidity, Working paper, Columbia and Wharton.

McFadden, Daniel, 1973, Conditional logit analysis of qualitative choice behavior, in P. Zarembka, ed., *Frontiers in Econometrics*, chapter 4, 105–142 (Academic Press: New York).

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, 2018, Commuting, migration, and local employment elasticities, *American Economic Review* 108, 3855–90.

Moretti, Enrico, 2013, Real wage inequality, *American Economic Journal: Applied Economics* 5, 65–103.

Nozawa, Yoshio, and Yancheng Qiu, 2020, Corporate bond market reactions to quantitative easing during the COVID-19 pandemic, Working paper, HKUST.

O'Hara, Maureen, and Xing Alex Zhou, 2021, The electronic evolution of corporate bond dealers, *Journal of Financial Economics* 140, 368–390.

O'Hara, Maureen, and Xing (Alex) Zhou, 2020, Anatomy of a liquidity crisis: Corporate bonds in the Covid-19 crisis, *Journal of Financial Economics* , Forthcoming.

Piazzesi, Monika, and Martin Schneider, 2016, Chapter 19 - housing and macroeconomics, volume 2 of *Handbook of Macroeconomics*, 1547–1640 (Elsevier).

Pintér, Gábor, Chaojun Wang, and Junyuan Zou, 2022, Size discount and size penalty: Trading costs in bond markets .

Quigley, John M, and Larry A Rosenthal, 2005, The effects of land use regulation on the price of housing: What do we know? what can we learn?, *Cityscape* 69–137.

Roback, Jennifer, 1982, Wages, rents, and the quality of life, *Journal of political Economy* 90, 1257–1278.

Rognlie, Matthew, 2015, Deciphering the fall and rise in the net capital share: Accumulation or scarcity?, *Brookings Papers on Economic Activity* 46, 1–69.

Rossi, Marco, 2014, Realized volatility, liquidity, and corporate yield spreads, *The Quarterly Journal of Finance* 4.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek, 2021, Ipums usa: Version 11.0 [dataset] .

Rust, John, 1987, Optimal replacement of gmc bus engines: An empirical model of harold zurcher, *Econometrica: Journal of the Econometric Society* 999–1033.

Saiz, Albert, 2010, The geographic determinants of housing supply, *The Quarterly Journal of Economics* 125, 1253–1296.

Schrimpf, Andreas, Hyun Song Shin, and Vladyslav Sushko, 2020, Leverage and margin spirals in fixed income markets during the covid-19 crisis, Working paper, BIS.

Schubert, Gregor, 2020, House price contagion and u.s. city migration networks, *Working Paper* .

Schultz, Paul, 2017, Inventory management by corporate bond dealers, Working paper, University of Notre Dame.

Scott, Paul, 2014, Dynamic discrete choice estimation of agricultural land use .

Sherwin, Rosen, 1979, Wages-based indexes of urban quality of life, in Peter Mieszkowski, and Mahlon Straszheim, eds., *Current Issues in Urban Economics*, 74–104 (Baltimore: John Hopkins University Press).

Üslü, Semih, 2019, Pricing and liquidity in decentralized asset markets, *Econometrica* Forthcoming.

Van Nieuwerburgh, Stijn, and Pierre-Olivier Weill, 2010, Why has house price dispersion gone up?, *The Review of Economic Studies* 77, 1567–1606.

Weill, Pierre-Olivier, 2007, Leaning against the wind, *Review of Economic Studies* 74, 1329–1354.