Title
Essays in Financial Economics

Author
Kim, Edward Taehoon

Publication Date
2023

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays in Financial Economics

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Management

by

Edward Taehoon Kim

2023

ABSTRACT OF THE DISSERTATION

Essays in Financial Economics

by

Edward Taehoon Kim
Doctor of Philosophy in Management
University of California, Los Angeles, 2023
Professor Andrea Lynn Eisfeldt, Chair

In Chapter 1, I explore the digital divide in homeownership outcomes as better financial opportunities increasingly move beyond the technology frontier. Low-income households derive significantly less savings from mortgage refinancing than their wealthy counterparts. I document that the rise of refinancing inequality in the United States can be partially explained by the gap in access to modern information and communications technology. Using granular spatial variation of a large-scale broadband subsidy program, I show that high-speed internet facilitates refinancing activity and reduces monthly mortgage payments. These effects are large and persistent, corresponding to a 5 percent increase in disposable income and up to $18,000 in total savings for low-income households. The growth of refinancing is pronounced in underserved areas with low access to bank branches and among populations that are likely to have low financial and digital literacy.

In Chapter 2 (with Andrea L. Eisfeldt and Dimitris Papanikolaou), we show that the recent underperformance of value investing strategies can be attributed to the mis-

measurement of intangible assets. We propose a simple improvement to the classic Fama and French value factor that incorporates intangibles and addresses differences in accounting practices across industries. Our intangible value factor prices assets as well as or better than the traditional value factor but yields substantially higher returns. This outperformance holds over the entire sample period, including in more recent decades during which value has underperformed. We also find evidence that the adjustment better identifies firms with superior fundamentals as measured by productivity, profitability, and financial soundness.

In Chapter 3 (with Marcelo Rezende), we study whether regulatory costs induce banks to search for yield by holding riskier assets. We test this hypothesis by analyzing a cost specifically imposed on the size and composition of a bank's balance sheet — deposit insurance premiums charged by the Federal Deposit Insurance Corporation (FDIC). Using supervisory data and a sharp cutoff in the schedule of deposit insurance premiums, we show that higher balance sheet costs indeed cause banks to substitute excess reserves (a liquid asset with no credit risk) for short-term interbank loans (a less liquid asset with credit risk). We argue that optimal deposit insurance pricing should account for this potential feedback effect.

The dissertation of Edward Taehoon Kim is approved.

Antoinette Schoar

Barney P. Hartman-Glaser

Mark J. Garmaise

Andrea Lynn Eisfeldt, Committee Chair

University of California, Los Angeles

2023

To Gina, Jaeyoung, Myungchul, and Rachel.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am deeply grateful to my advisor, Andrea Eisfeldt, for her unwavering support and intellectual guidance. Committee members Mark Garmaise, Barney Hartman-Glaser, and Antoinette Schoar also provided valuable mentorship. I thank UCLA faculty Mikhail Chernov, Stuart Gabriel, Mark Grinblatt, Valentin Haddad, Bernard Herskovic, Shohini Kundu, Lars Lochstoer, Francis Longstaff, William Mann, Tyler Muir, Stavros Panageas, Gregor Schubert, Avanidhar Subrahmanyam, Xinxin Wang, Ivo Welch, David Wessels, and Jinyuan Zhang, for shaping my knowledge of the discipline, offering constructive advice for my research, and inviting me to serve a fulfilling role in the lecture hall.

For helpful feedback and discussions, I thank seminar participants at UCLA Anderson, Michigan Ross, USC Marshall, Indiana Kelley, Penn State Smeal, Rochester Simon, Notre Dame Mendoza, Virginia Tech Pamplin, Federal Reserve Board, U.S. Department of the Treasury (Office of Financial Research), the FDIC, and the Society of Financial Studies Cavalcade, as well as editors, anonymous referees, and discussants. Chapter 2 is a version of Andrea L. Eisfeldt, Edward T. Kim, Dimitris Papanikolaou; Intangible Value, *Critical Finance Review*, 2022, 11(2): pp. 299-332, `http://doi.org/10.1561/104.00000113`. Chapter 3 is a version of Edward T. Kim, Marcelo Rezende; Deposit Insurance Premiums and Bank Risk, *The Review of Corporate Finance Studies*, 2023, 12(2): pp. 291-325, `https://doi.org/10.1093/rcfs/cfad004`.

I am grateful for my former colleagues at the Federal Reserve Board, especially my supervisors John Driscoll and Rebecca Zarutskie. I also thank Dave Donaldson for introducing me to academic research and Bengt Holmstrom for providing the original inspiration towards pursuing this degree.

I had the pleasure of working closely with fellow Ph.D. students Salil Gadgil, Paul

Huebner, Denis Mokanov, Paul Momtaz, James O'Neill, Gabriel Cuevas Rodriguez, Zachary Sauers, Clinton Tepper, Danyu Zhang, and recent alumni Darren Aiello, Mahyar Kargar, and Yu Shi. I thank UCLA Anderson's current and former staff for their assistance throughout the doctoral program and acknowledge the generous financial support of UCLA Graduate Division, UCLA Anderson, UCLA Fink Center for Finance, and UCLA Ziman Center Gilbert Program in Real Estate, Finance, and Urban Economics.

VITA

2015	B.S., Economics, Massachusetts Institute of Technology, Cambridge, MA
2015–2017	Research Assistant, Federal Reserve Board, Washington, DC

CHAPTER 1

The Digital Divide and Refinancing Inequality

Mortgage refinancing is an important mechanism for household wealth accumulation in the United States; however, many Americans do not refinance their mortgages optimally due to frictions such as high origination costs or limited financial sophistication (Campbell, 2006). This phenomenon is concentrated among low-income and minority households, implying a potential imbalance in the transmission of monetary policy during economic downturns that can exacerbate wealth inequality. In this paper, I study whether at-home access to modern information and communications technology can help mitigate refinancing frictions. Specifically, I demonstrate that access to broadband internet increases refinancing activity and reduces housing costs for low-income households.

High-speed internet can significantly reduce the shadow costs associated with applying to refinance a mortgage. Using the internet, an applicant can easily exchange paperwork by e-mail, link financial accounts online to expedite credit verification, and spend less time meeting with a loan officer or visiting a bank branch. Indeed, processing times for mortgage applications at online lenders are estimated to be 15 to 30 percent shorter than at their physical counterparts, with a larger effect for refinance loans (Fuster et al., 2019). To the extent that online resources allow households to obtain information about the value of refinancing, the internet can also reduce the incidence of suboptimal refinancing driven by behavioral mistakes.

Despite the internet's large role in streamlining the refinance process, it is inaccessible to millions of American households living without a wired broadband connection at home. The persistent gap in access to information technology, known as the "digital divide," has become an important policy issue in recent decades due to its influence on household well-being (White House, 2022). In 2019, less than 70 percent of the population reported having a broadband subscription at home, with low-income households reporting significantly lower subscription rates (Figure 1.1). This trend is not entirely driven by the lack of physical access to a broadband provider; of the low-income households living in urban areas with near-complete broadband coverage, only 65 percent subscribed to broadband during this period.

Studying the effects of broadband access on refinancing inequality is difficult for several reasons. First, the spatial distribution of broadband providers is correlated with subscriber characteristics such as employment and educational attainment. As these characteristics are also correlated with refinancing demand, estimates of refinancing outcomes that relate to heterogeneity in broadband availability will most likely be biased. Second, it is difficult to observe exogenous changes in broadband adoption by households, especially for low-income homeowners that tend to refinance suboptimally. As a result, little is known about the extent to which broadband access can reduce refinancing frictions.

To address these empirical challenges in quantifying the effect of broadband access on refinancing, I analyze the Internet Essentials program by Comcast, one of the largest broadband providers in the United States. Introduced in 2012 to receive regulatory approval for a merger, Internet Essentials heavily subsidized broadband subscription fees to qualifying low-income households. The monthly cost of $9.95 was up to 75 percent lower than a that of a comparable regular plan, and all fees related to activation and equipment (averaging more than $100 upfront and up to $10 per month, respectively) were waived. The program became highly successful, connecting 750,000 American families (or 3 million individuals) nationwide in the first five years

(Comcast Corporation, 2016). Internet Essentials is a suitable setting to study refinancing behavior due to its unique properties. First, it was immediately available in all of Comcast's existing service areas. This method of rollout is important for identification because physical infrastructure expansions associated with other broadband initiatives not only take time but also can increase local house prices, confounding the estimated impact of broadband access on refinancing (Knutson, 2015). Second, Internet Essentials was directly aimed at increasing broadband take-up by low-income households making less than around $40,000 per year — the group that exhibits low refinancing behavior most prominently. Third, internet usage at broadband speeds would have been a binding constraint for households to access banking services during the study period. Lastly, the program coincides with the prolonged recovery period after the Great Recession when refinancing incentives and potential monetary savings were high throughout the income distribution.

This paper exploits geographic, temporal, and household-level variation in Internet Essentials eligibility to estimate the impact of broadband access on refinancing demand and mortgage costs. Specifically, I compare the outcomes of eligible and ineligible low-income households across census tracts with and without Comcast service before and after 2012. The identifying assumption is that within-census tract differences in refinancing outcomes between eligible and ineligible households are uncorrelated with Comcast coverage except through the introduction of the Internet Essentials program. Indeed, I do not find any violation of the common trends assumption under this empirical setting. I construct a unique data set that matches Comcast coverage rates at the census tract level to the universe of refinance applications and originations by income eligibility between 2008 and 2015. I also enhance my analysis using a matched panel data set of prepayment propensities for home purchase mortgages originated between 2004 and 2008, as well as American Community Survey (ACS) microdata on mortgage payment burdens.

I find that improved broadband access leads to a strongly positive impact on refinancing outcomes. In particular, both the number of submitted applications and originated loans increased by 6 percent as a result of Internet Essentials. Importantly, household financial gains are driven by behavioral changes along the extensive margin (increased likelihood of refinancing) and not through differential effects along the intensive margin (lower interest rates). Using household-level survey data, I corroborate the findings of increased refinancing propensity with evidence of decreased mortgage payment burdens. In addition, I show that the results are in large part driven by census tracts with limited access to physical bank branches, implying that broadband promotes access to financial services for the underbanked. Treatment effects are also stronger for households with low educational attainment, which suggests a digital and financial literacy channel for refinancing.

The economic magnitudes of these results are significant: the average low-income household that refinanced its mortgage between 2012 and 2015 would have saved up to \$100 per month on mortgage payments even after accounting for the nominal cost of subscribing to Internet Essentials. This translates to a 5 percent increase in monthly disposable income and total household wealth gains of up to \$18,000 in present value terms, which accounts for about 10 percent of the average net worth of homeowners in this income bracket. I estimate that the program generated up to \$100 million in additional refinance savings across Comcast area households and reduced refinancing inequality by up to 14 percent.

These empirical findings are robust to several validation and falsification tests. To start, I verify that the results hold when using mortgage prepayment as an alternative measure of refinancing. Second, I assign placebo treatment indicators for AT&T and Charter coverage instead (the next two largest broadband providers by subscriber count) and find no effects of broadband access on refinancing outcomes. Third, the results disappear when I use households with incomes marginally above the eligibility threshold as the treated group, supporting the identifying assumption that the program only affected eligible low-income households. Fourth, treatment effects are concentrated in census tracts with a high likelihood of being affected by Internet Essentials.

Related Literature. This paper is related to the growing literature on the determinants of mortgage refinancing behavior. Campbell (2006) documents low levels of refinancing among low-income borrowers in the early 2000s. In more recent work, Andersen et al. (2020), Agarwal et al. (2013, 2016, 2020), Defusco and Mondragon (2020), Gerardi et al. (2020), Gerardi et al. (2021), Goodstein (2013), Johnson et al. (2018), Keys et al. (2016) all find evidence of suboptimal refinancing behavior driven by income and race, particularly during the aftermath of the Great Recession and the recent COVID-19 pandemic. Other works identify specific behavioral channels such as financial illiteracy (Agarwal et al., 2017; Bajo and Barbi, 2018), inattention (Byrne et al., 2022), distrust of financial institutions (Johnson et al., 2018; Yang, 2021), and peer effects (Maturana and Nickerson, 2018). To my best knowledge, this paper is the first to analyze the role of a relatively understudied but influential aspect of everyday life — broadband internet — that can impact both the demand for and supply of refinance credit especially for disadvantaged populations. My results are also relevant for the implementation of broadband infrastructure initiatives, which has become an integral part of public policy discourse in recent years.

The literature on the role of financial technology in household finance, most notably Philippon (2016), Buchak et al. (2018), Di Maggio et al. (2021), and Bartlett et al. (2022), has highlighted technology's large impact on mortgage market composition and lending practices. This paper serves as a complement to Fuster et al. (2019), who document the large role fintech lenders play in reducing processing times for mortgage applications submitted online. Importantly, the authors find no effect of broadband access on mortgage outcomes using the rollout of Google Fiber as an instrument. By studying a national program that did not require low-income customers to pay large upfront costs, I provide suggestive evidence that broadband internet can indeed reduce refinancing frictions. In addition, recent works on financial inclusion highlight the persistent importance of bank branches in the modern era (Brown et al., 2019; Célerier and Matray, 2019; Fonseca and Matray, 2022; Jung and Zentefis, 2022) and the implications of digital disruption (Jiang et al., 2022). Yogo et al. (2021) also find that financial participation depends on household income rather than race or access to financial services. My paper contributes to this literature by showing that the inability of low-income households to afford broadband internet can be a significant impediment to financial inclusion.

Outline. The remainder of the paper is structured as follows. Section 1.1 describes the institutional background on mortgage refinancing, broadband access, and the Internet Essentials program. Section 1.2 describes the data and empirical methodology. Section 1.3 discusses the main results and studies the relevant mechanisms. Section 1.4 provides robustness checks as well as falsification tests. Section 1.5 concludes.

1.1 Background

1.1.1 Mortgage Refinancing

Households use mortgages to purchase a new property or refinance an existing mortgage on a previously purchased property. Since most mortgages in the United States are fixed-rate loans without prepayment penalties, a refinance allows households to reduce their cost of credit when interest rates fall. In essense, the refinance decision is a call option that should be exercised when the original loan is "in the money" after adjusting for interest rate differentials and closing costs. Refinancing constitutes a large segment of residential real estate markets, accounting for more than half of all mortgage originations by volume between 2005 and 2015 (Haughwout et al., 2021).

Homeownership is the primary source of wealth creation among American families, with about 65 percent of the population residing in owner-occupied units as of 2019. Understanding what drives households to refinance their mortgage is important in light of the weight placed on homeownership in their portfolios, representing between 30 and 40 percent of household net worth (Current Population Reports, 2019). As such, refinancing to lower mortgage payments is one of the most consequential decisions a household makes throughout its lifetime. The importance of housing is particularly large for low-income households, whose homes account for over 80 percent of their total wealth. I first document the prevalence of homeownership among low-income households. According to the National Association of Realtors, around 38 percent of low-income households resided in owner-occupied units in 2010. This group's contribution to the housing market is not trivial; households with annual income less than $35,000 purchased home mortgages worth $780 billion between 2001 and 2008, with an average home value at origination of $120,000 and monthly payments of $700 over 30 years. Housing cost burdens are also disproportionately large for this income group, with more than half of homeowners paying 30 percent or more of their monthly disposable income on housing. Reducing mortgage payments through refinancing, therefore, is an important way to increase household net worth through additional savings.

Prior research has documented that many households fail to refinance their mortgages when it is optimal to do so (Agarwal et al., 2016; Keys et al., 2016; Johnson et al., 2018; Andersen et al., 2020). These financial mistakes are particularly pronounced among low-income households; of the mortgages originated between 2004 and 2008 by households making less than $35,000 in annual income, only around 65 percent were refinanced at any point between 2009 and 2015, the period during which mortgage interest rates fell by an average of 1.5 to 2 percent. This stands in stark contrast to the refinancing propensity of loans originated by households making more than $75,000 (80 percent). This trend is monotonic throughout the income distribution and also prevalent in large central metro areas, which tend to have more resilient banking systems (Figure 1.2). The pronounced errors at the lower end of the income distribution persists even after controlling for predictors of financial distress during the Great Recession, such as debt-to-income ratio (DTI), loan-to-value ratio (LTV), and credit score. This paper provides evidence that borrower frictions relating to information technology plays an important role in explaining these disparities.

1.1.2 Broadband Internet in the United States

Broadband technology, which grew in prevalence since the early 2000s, allows households to use the internet for all aspects of life including work, education, and entertainment. In this paper, I define broadband as a residential, high-speed, wireline internet service available in a given geographic area. I focus on residential (as opposed to commercial) service as it is relevant to at-home household financial decisions. High-speed status is determined by whether a service meets the standards for broadband set by the Federal Communications Commission (FCC). The minimum download speed for broadband was 4 megabits per second (Mbps) during the study period, which is adequate for general web browsing, e-mail communication, and some video streaming at low bandwidths.[1] The predecessor technology of dial-up internet, on the other hand, typically has a maximum download speed of 56 kilobytes per second (Kbps), or 1.4 percent of the speed of broadband internet. Dial-up internet is not considered in this paper as the technology has struggled to keep up with the increasingly complex needs of everyday internet usage. Lastly, I only consider wireline service provided through physical broadband infrastructure. This is because wireless

[1] The 4 Mbps minimum speed standard for broadband was set in 2010 and then revised up to 25 Mbps in 2015.

networks accessed through mobile devices were not reliable or advanced enough to replace broadband during the late 2000s and early 2010s.

The lack of broadband internet at home, particularly in urban areas, can largely be attributed to low affordability. Figure 1.3 shows a clear negative relationship between census tract poverty rates and broadband subscription rates. This trend is not driven by limited access to a broadband provider. In fact, more than 90 percent of the urban population in the United States lived in areas with broadband service by 2015, while only 70 percent (60 percent for low-income groups) reported actually having a broadband subscription.[2] Survey results from the Pew Research Center reveal that the price of subscription (59 percent) and cost of computer equipment (45 percent) are the top two reasons for not subscribing to broadband (Horrigan and Duggan, 2015). While the urban-rural disparity in broadband coverage is an important access-driven cause for the digital divide, I focus on cost-driven disparities in subscription conditional on having access to infrastructure. This framework is useful for identification because it is invariant to unobservable differences in broadband service quality and customer demand across urban and rural areas.

1.1.3 Broadband and Refinancing Inequality

At-home internet access is relevant for refinancing inequality due to the unique properties of a refinance mortgage. First, refinancing is largely standardized and compatible with technological innovation. In most interest rate refinances, the housing asset in question is already determined and the prospective borrower is in good standing on the existing mortgage.[3] Borrower uncertainty is thus low, allowing the refinance

[2] Statistics are compiled from the 2015 FCC Broadband Progress Report and author's calculations using ACS 2017 5-year estimates.

[3] Since a refinance requires current homeownership, it is not determined by exogenous motives to move into or out of a dwelling. This is important as it allows the borrower pool to be invariant from significant income shocks or migrational incentives.

process to be streamlined and automated. Recent innovations in online approval and underwriting technology have led to a notable decrease (up to 30 percent from an average of 51 days) in processing time for refinance applications (Fuster et al., 2019). The internet has also enabled both bank and non-bank lenders to reach populations outside their immediate geographic markets, improving the access to refinancing credit for underbanked households.

Second, refinancing involves high shadow costs for borrowers (i.e., time and cognitive effort) that can be drastically reduced through internet usage. A refinance typically takes several months to complete, primarily due to stringent documentation requirements that include recent pay stubs, tax returns, W-2s, homeowners insurance policies, asset statements (e.g., checking, savings and investment) and debt statements (e.g., credit card and automobile). For the majority of American households that use online banking, these materials can be conveniently accessed and transmitted online with a computer and broadband connection.[4] Furthermore, applicants with broadband can use e-mail to communicate with a loan officer and make fewer branch visits. To the extent that the internet can also increase households' awareness and provide resources to shop around for lenders and rates, broadband access at home has become an important way to reduce the shadow costs associated with refinancing. Indeed, Figure 1.4 shows that local area broadband access is correlated with online search activity for information about refinancing and current mortgage rates.

In this paper, I argue that refinancing inequality arises in part due to heterogeneity in broadband access. As low-income households typically face volatile employment prospects and work longer hours, they may find it particularly difficult to fulfill the verification and qualification requirements for a refinance without at-home internet. Moreover, these households tend to be underbanked and are less confident in their

[4] 55.1 percent of the population reported using online banking and one-third reported using it as the main method to access bank accounts (Federal Deposit Insurance Corporation, 2013).

ability to get approved for other types of credit, suggesting that both access to and demand for financial services via brick-and-mortar branch networks is limited.[5] Lastly, information frictions regarding upfront costs (which can be rolled into payments or entirely waived via government programs) can further reduce refinancing activity for low-income households that typically lack savings in financial assets.[6] Figure 1.5 shows that broadband access is correlated with disparities in realized refinancing outcomes: voluntary prepayment propensities for households with high refinance likelihood are generally lower in census tracts with limited broadband access, with a larger gap for the bottom income decile.

1.1.4 Internet Essentials Program by Comcast

Internet Essentials by Comcast provides a useful quasi-experimental setting to study the digital divide in mortgage refinancing. Comcast is one of the nation's largest internet service providers (ISPs), operating in 39 states and the District of Columbia and covering 48 million households at the time of the study. Internet Essentials was originally conceived to garner the FCC's support for a proposed merger with NBC Universal, a media and entertainment conglomerate corporation. The FCC ultimately approved the merger and enforced Comcast's commitment to institute the low-income subsidy program to promote public interest (FCC, 2012). In the beginning of 2012, Internet Essentials was made available in all Comcast coverage areas nationwide and became the first comprehensive program of its kind by a major ISP.

[5] 27 percent of households with less than $40,000 in annual income were underbanked, compared to 11 percent for households with income above $100,000. 32 percent of low-income respondents reported not being confident in their ability to be approved for a credit card loan, compared to 7.2 percent for high-income respondents (Report on the Economic Well-Being of U.S. Households in 2015).

[6] Bhutta and Dettling (2018) find that only 51 percent of households in the bottom income quartile had at least $400 in savings for an unexpected expense, and 17 percent reported having savings worth 3 months of expenses.

In an effort to achieve the FCC's mandate of fostering competition and benefiting consumers through reasonably priced broadband offerings, Internet Essentials significantly reduced the cost of broadband subscription. Enrolled households received high-speed broadband (15 Mbps download and 2 Mbps upload) for a $9.95 monthly fee plus applicable taxes, which is about 75 percent lower than the average cost of a comparable unsubsidized broadband plan (Hussain et al., 2013). Moreover, all one-time installation and activation fees (up to $100) as well as modem and router rental fees (up to $20 per month) were waived. Fee savings over a three year period would have exceeded $1,720, which is a sizeable amount for eligible households with an average annual income of $30,000. Internet Essentials also offered subsidized computers for $149.99 and provided digital literacy training resources through online offerings as well as an extensive network of over 9,000 community organizations, libraries, and elected officials.

Eligibility requirements for Internet Essentials were carefully designed to maximize impact and administrative convenience. First, a household must reside in an area that is served by Comcast at the time of application. Second, a household qualifies if it has a child receiving free or reduced-price lunch under the National School Lunch Program (NSLP). These meal benefits in turn depend on household size and income. Specifically, eligibility is restricted to households with annual income below 185 percent of the federal poverty limit (FPL), which translates to around $35,000 for a three-person family and $42,000 for a four-person family during the study period.[7] Third, an applicant must not have any past-due debt to Comcast and cannot have been a Comcast subscriber in the preceding 90 days. This restriction, along with the high concentration and visibility of Comcast as the major ISP in most of its coverage areas, makes it likely that new subscribers did not have an existing broadband

[7] In 2010, 31.8 million children participated in the NSLP nationwide (U.S. Department of Agriculture, 2019).

subscription. Indeed, 80 percent of Internet Essentials customers reported not having any broadband internet service at some point in the past (Comcast Corporation, 2016). Internet Essentials was principally rolled out through extensive public service announcement campaigns as well as partnerships with thousands of school districts, non-profit organizations, and city councils. Comcast also streamlined the application process in the early years by auto-approving households with children attending majority low-income schools.

Internet Essentials was highly successful, connecting more than 750,000 low-income families (or 3 million individuals) between 2012 and 2016. Importantly, the program grew in urban areas more quickly due to the strong emphasis on community partnerships; 75 percent of the subscribers in the first five years came from 10 of the 40 states and the top 10 cities accounted for 25 percent of subscriptions in this period (Comcast Corporation, 2016). Internet Essentials rapidly became an integral part of everyday life for low-income households, with 89 percent of subscribers reporting using the internet almost every day. Table 1.1 reports the average characteristics of subscribers and statistics on internet usage. A large fraction of Internet Essentials subscribers are represented by racial minorities (black or hispanic) with low income and low educational attainment. In terms of common internet usage other than children's schoolwork, a majority of subscribers reported using the internet to find general information (92 percent), access e-mail (80 percent), and connect with others on social media (71 percent). Importantly, 65 percent of subscribers said that banks or other financial institutions expect them to have internet access at home. In a subsequent survey, 42 percent reported using the internet to access banking and financial services (Horrigan, 2014, 2019).

1.2 Methods and Data Description

1.2.1 Empirical Design

I discuss two important challenges for quantifying the causal effect of Internet Essentials on refinancing. First, it is difficult to compare refinance outcomes of income-eligible (treated) and income-ineligible (control) households within Comcast areas due to non-parallel trends. As income is a primary predictor of mortgage principal, and by extension, monetary savings from refinancing, ineligible households with marginally higher incomes are more likely to refinance early when interest rates fall.[8] Moreover, refinancing is typically a one-shot decision for most homeowners due to large origination costs. This leads to a natural attrition of the ineligible group's potential refinance pool in the early years following the Great Recession. Thus, any positive effects of Internet Essentials' introduction in 2012 will be biased upwards by the increasing trend of refinancing activity by eligible low-income households throughout the recovery period.

Second, it is not feasible to directly compare refinance outcomes of eligible households in Comcast and non-Comcast areas. Importantly, Comcast has near-complete coverage in certain major cities (e.g., Chicago, Sacramento, Miami, Houston) and is entirely absent in others (e.g., Los Angeles, New York, Dallas), making it difficult to identify two regions within a small geographic footprint with varying levels of coverage. As a result, a standard study of differences in refinancing behavior between Los Angeles and Sacramento (or between Chicago and New York) is likely to be driven by unobservable confounders. Even after controlling for economic and financial indicators that motivate a household's refinance decision (for instance, house prices and interest rates), I cannot rule out the impact of factors such as industry-by-tract

[8] This fact is further supported by the monotonic increase in refinancing propensities visualized in Figure 1.2.

employment outcomes, migration patterns, or nuanced changes in lending standards that may bias the estimates.

To overcome these limitations, I study Internet Essentials' impact on refinancing by addressing both the variation in geographic coverage and income eligibility, in conjunction with temporal variation pre- and post-program launch. In particular, I use a difference in difference in differences ("triple differences") design introduced by Gruber (1994) to compare changes in the *gap* of refinancing outcomes between eligible and ineligible groups across Comcast and no Comcast census tracts. Under my empirical setting, any confounders at the census tract level that impact both eligible and ineligible groups concurrently will be absorbed. Identification relies on the assumption that the difference in outcomes between the two eligibility groups within a census tract will not vary with Comcast coverage before and after 2012, except through the impact of Internet Essentials.

Figure 1.6 illustrates the intuition behind the triple differences design. All three panels plot the residualized number of annual refinance originations by each eligibility group at the census tract level — one of the main outcome variables of interest. The top panel shows that eligible and ineligible groups within Comcast areas follow divergent trends in refinancing behavior prior to the program's launch in 2012. In the middle panel, I show that the two groups in no Comcast census tracts also exhibit similar trends in refinancing behavior throughout the study period. Lastly, the gap in refinancing originations between eligible and ineligible groups in non Comcast areas is consistent with the corresponding gap in Comcast areas leading up to the program's introduction in 2012 (bottom panel).

1.2.2 Data Sources

Comcast Coverage Rates. I compute coverage rates for Comcast and other major ISPs using service availability data obtained from the National Telecommunications and Information Administration (NTIA)'s State Broadband Initiative.[9] As required by law, each ISP self-reports whether it offered any type of internet service in a given census block on a biannual basis. I restrict the provider responses to those that can be classified as broadband service and aggregate the information up to the census tract level to compute coverage rates.

Mortgage Applications and Originations. The Home Mortgage Disclosure Act (HMDA) provides loan-level data on the near-universe of mortgage applications in the United States. To standardize the borrower pool and minimize the effect of refinancing incentives driven by exogenous factors, I restrict the sample to owner-occupied, one- to four-family, conventional refinance mortgages.[10] Importantly, HMDA data reports an applicant's income and location of the property at the census tract level, along with demographic characteristics such as race and sex. The main dependent variable in my analysis captures changes in refinancing demand and outcomes over time. For each year between 2008 and 2015, I count the number of refinance applications submitted by eligible and ineligible households in a given urban census tract. I additionally tally the number of originated mortgages and compute denial rates for each eligibility group by taking the ratio of denials to total applications.

Prepayment Activity and Loan-Level Covariates. Prepayment refers to the payment of a mortgage's principal before maturity. While there may be many reasons for prepayment (including foreclosure), I focus on voluntary prepayment as an additional proxy for refinancing activity.[11] First, I measure prepayment of mort-

---

[9] Recent provider data after 2014 are compiled centrally by the FCC through Form 477. The FCC also reports census tract- and county-level information on the number of broadband connections per 1,000 households.

[10] Conventional mortgages are not insured or guaranteed by the Federal Housing Administration (FHA), Veterans Administration (VA), Farm Service Agency (FSA) and Rural Housing Service (RHS).

[11] The vast majority of voluntary prepayments are as a result of refinancing, and prior research has studied prepayment speeds as a proxy for refinancing activity (Schwartz and Torous, 1989; Stanton, 1995; Longstaff, 2005; Deng and Quigley, 2012).

gages originated between 2004 and 2008 using loan performance data supplied by two major government sponsored enterprises (GSEs). In particular, I assign an indicator for whether a 30-year fixed rate mortgage purchased by Fannie Mae or Freddie Mac is prepaid between 2008 and 2011 (pre-Internet Essentials), and another indicator for whether the mortgage is prepaid between 2012 and 2015 (post-Internet Essentials). While the performance data also contain the location of the home at the 3-digit zip code level as well as useful loan characteristics, they importantly do not report borrower income that is required for assigning treatment status. Thus, I programmatically merge the GSE filings to HMDA data using six exact match categories (year of origination, agency, owner occupancy, loan type, number of applicants, and loan amount) and a fuzzy match category (location).[12] The resulting data set covers between 20 and 30 percent of all mortgages originated and sold to the two GSEs. In addition to the demographic characteristics available in HMDA, the matched data provides important loan-level covariates at origination such as interest rates, debt-to-income ratios (DTI), combined loan-to-value ratios (CLTV), and credit scores. I also calculate a time-varying measure of each loan's remaining maturity at the time prepayment is observed.

Interest Rates. I test whether broadband access reduces the incidence of sub-optimal refinancing by analyzing interest rate outcomes. Loan-level interest rates are available in the GSE performance data, while borrower income is only reported in HMDA. I employ the matching process detailed above to merge the two data sources for refinance mortgages originated between 2008 and 2015. Specifically, I obtain interest rates for a representative subset of owner-occupied, one- to four-family, conventional 30-year mortgages sold to Fannie Mae and Freddie Mac.

Mortgage and Rental Costs. I collect information on households' mortgage and rental payments from the Integrated Public Use Microdata Series (IPUMS) of the

[12] Further details on the matching process can be found in the Internet Appendix.

American Community Survey (ACS) 1-year estimates. De-identified microdata are published for all survey respondents each year. The survey reports mortgage or rental payments made by each household in dollar amounts as well as relevant covariates on home value and demographic information (age, gender, race, educational attainment, etc.). Importantly, the questionnaire contains details about income and household composition that help refine the assignment to Internet Essentials eligibility. Geographic location is identified at the PUMA level (average population above 100,000), which is significantly larger than a census tract (average population of 4,000).

House Prices and Average Income. In my main empirical analysis, census tract level trends in house prices and homeowner income are absorbed by year fixed effects. While low-income treatment and control groups are likely to experience shocks in these factors concurrently, I additionally incorporate controls for group-level changes in economic outcomes using HMDA data. In particular, I construct a time-varying proxy for house prices as the logarithm of average originated loan amounts by eligibility group. Similarly, the logarithm of average income measures changes in income levels among borrowers in each group. For specifications that do not rely on within-tract variation in house prices over time, I use annual house price index (HPI) data published by the Federal Housing Finance Agency (FHFA). The data is available at the census tract level and capture the evolution of overall refinancing incentives for homeowners.

Bank Branch Access. I compile location information for bank branches using data from the Federal Deposit Insurance Corporation (FDIC)'s Summary of Deposits. The data includes precise geographic coordinates for all FDIC-insured financial institutions each year. For each census tract, I compute the number of full service ("Brick and Mortar" or "Retail") bank branches that are within a 2 mile radius of the population centroid as of 2010. Location information for the center of population is obtained from the Census.[13]

Fintech Lenders. Banks and financial institutions that allow a customer to complete the entire mortgage origination process online are classified as fintech lenders. I use the definition of fintech lenders suggested by Buchak et al. (2018) and Fuster et al. (2019). I then match these fintech classifications to HMDA data using the respondent identifier associated with each mortgage application.

Other Demographics. Broadband and refinancing inequality are crucially driven by disparities in economic outcomes across urban and rural areas. To address this, I classify census tracts into urban and rural areas using the scheme provided by the National Center for Health Statistics (NCHS).[14] In particular, I use the 2006 delineation of county-level urbanicity and match it to each census tract. Demographic characteristics such as tract-level unemployment, broadband usage, and educational attainment, are obtained from the ACS summary and microdata files.

1.2.3 Comcast Coverage Rates and Income Eligibility

Assignment to treatment in my empirical setting relies on two important sources of variation: Comcast coverage rates and income eligibility. To calculate Comcast coverage rates, I first restrict the NTIA's block-level provider data to connection types that qualify as broadband according to the definition used in this paper. As census blocks are a clean subset of a census tract, I then aggregate the block-level

[13] While most studies on "banking deserts" measure branch access within a 10-mile radius of the population centroid. I follow the 2-mile radius convention used by Covas (2019). As the census tracts in my sample are geographically small (about 7 square miles on average) and concentrated in urban clusters, using the measure using the 10-mile radius is likely to overstate true bank access.

[14] https://www.cdc.gov/nchs/data_access/urban_rural.htm.

data as of December 2011 (the year prior to Internet Essentials) by calculating:

$$Comcast_{c,2011} = \frac{\sum_{b=1}^{c} Population_{b,2010} \times \mathbf{1}(Comcast_{b,2011})}{Population_{b,2010}}, \tag{1.1}$$

where $Population_{b,2010}$ refers to the population of block b and $\mathbf{1}(Comcast_{b,2011})$ is an indicator for whether Comcast provides broadband service in block b in 2011. $Comcast_{c,2011}$ captures the fraction of tract c's population that has access to Comcast broadband.[15] I address possible time-varying changes in coverage by using the same method to calculate $Comcast_{c,2014}$ and taking the average of the two rates to compute $Comcast_c$. Panel (a) in 1.7 presents a histogram of $Comcast_c$ in large central metropolitan counties, which exhibits a clear bimodal distribution with peaks at 0 and 100 percent. This distribution enables clean identification of treated census tracts that have near-complete Comcast coverage and control census tracts with no Comcast presence. For placebo tests, I use the same methodology to construct coverage rates for AT&T and Charter, the next two largest ISPs by subscriber count.

Eligibility for Internet Essentials also depends on whether a household has at least one child that receives free or reduced-price lunch at school. The baseline criteria for lunch benefits is in turn determined by low-income status given the size of the household, neither of which I can directly observe from the HMDA or GSE data. In my analysis, I first assume that all homeowners have a school-aged child between ages 6 and 18. Next, I assign low-income status based on a four-person household, which corresponds to the average household size of Internet Essentials subscribers. The income threshold for a four-person household increases slightly each year to account for inflation and averages $42,000 between 2008 and 2015. I classify all households

[15] Under NTIA's reporting requirements, a provider can report an entire census block as "served" if a single household can be connected to service on demand. As blocks cover a small geographic footprint in urban metropolitan areas, the study's setting is less likely to suffer from overestimation bias of actual broadband access.

with income less than 185 percent of the FPL for a three-person household ($35,000) as eligible and households with income more than 185 percent of the FPL for a five-person household ($49,000) as ineligible. Households with income between the three- and five-person household thresholds are excluded from analysis to account for possible measurement error. This classification method allows me to compute an intent-to-treat effect that is plausible as long as I can rule out differential biases in assignment across geographic areas that correlate with Comcast coverage. Finally, I further restrict the control group to households with income below 185 percent of the FPL for a six-person family ($57,000). This upper bound allows me to focus on two groups with relatively similar income. The resulting annual thresholds for Internet Essentials eligibility are tabulated in Table 1.2.

For analyses using ACS data, I directly observe income, family size, and the existence and age of children at the household level. The data thus allows a cleaner assignment to Internet Essentials eligibility. In particular, I classify treated households as those with at least one school-aged child and with income less than 170 percent of the FPL based on actual household size. Control households either have incomes between 200 and 270 percent of the FPL, do not have a school-aged child, or both. Again, I drop all households making more than 270 percent of the FPL for comparability as well as households with income between 170 and 200 percent of the threshold to address measurement error. In addition, I construct an alternative control group with the same income levels as the treated group (below 170 percent of FPL) but without a school-aged child. This final classification enables the most direct analysis of households that share similar economic characteristics but differ in eligibility.

1.2.4 Final Sample

I restrict my sample to census tracts in large central metropolitan counties as defined by the NCHS. This step is relevant because Internet Essentials' initial success was primarily led by Comcast's partnerships with local governments and school districts in urban areas. Limiting the analysis to urban areas thus guarantees the highest likelihood of broadband subscription by eligible low-income households in the years following the program's launch. I also drop census tracts that did not receive any refinance applications (regardless of income) in any given year between 2008 and 2015.

The final sample consists of 5,256 census tracts covering 57 MSAs. 2,430 tracts have higher than 50 percent Comcast coverage and 2,826 have less than 50 percent coverage.[16] Table 1.3 reports 15 high Comcast and 15 no Comcast metropolitan statistical areas (MSA) ranked by population served. The lack of overlap between the two groups implies that Comcast does not operate alongside other major ISPs in cities and rules out potential spillover effects across adjacent tracts with opposite coverage status. Additionally, the large number of census tracts within each MSA provides support for an empirical strategy that controls for tract-specific trends.

In Figure 1.7, panel (b), I map all census tracts in my sample and show that Comcast coverage also does not exhibit any patterns of regional clustering. Importantly, most of the census tracts without Comcast have permanent presence of either AT&T or Charter. This means that broadband environments in Comcast and no Comcast areas will mostly be similar; both areas will have comparable levels of broadband provider access, network quality and customer service, with the only major difference being that eligible households in Comcast census tracts could save up to 75% on their subscription costs starting in 2012.

[16] I use a continuous measure of Comcast coverage as the treatment indicator in all regression analyses. This is largely inconsequential because the distribution of coverage rates, as shown in Figure 1.7, is highly concentrated at 0 and 100 percent. All results are robust to using an indicator variable for whether $Comcast_c$ is above 70 percent (treated) and below 30 percent (control).

Table 1.4 presents descriptive statistics for select variables in Comcast and no Comcast census tracts. While Comcast census tracts are slightly less populated on average, the two groups share very similar characteristics in terms of income distribution, urbanicity, median age, average household size, owner-occupancy rates, mortgage cost burdens, employment rates, and education levels. Interestingly, Comcast census tracts tend to have a higher concentration of bank branches near the population center, and also exhibit higher broadband subscription rates.

Table 1.5 further reports descriptive statistics for mortgages and homeowner demographics in Comcast and no Comcast census tracts by eligibility status. Columns 2 and 3 (5 and 6) show that ineligible households have higher income and credit scores, purchase higher-valued homes, and receive more favorable interest rates than their eligible counterparts. Note that even control households still have substantially lower income relative to the rest of the population (Columns 1 and 4). For the average low-income mortgage originated between 2004 and 2008, the interest rate differential for refinancing between 2008 and 2011 was between 1.2 and 1.3 percentage points, which exceeds the typical threshold for optimal refinancing cited in the literature (Agarwal et al., 2013). Average interest rates fell further by a percentage point between 2012 and 2015, which contributed to a large refinancing wave throughout the income distribution. Comcast census tracts also tend to have a larger fraction of black homeowners and smaller fraction of hispanic homeowners than low Comcast census tracts. In general, the difference in observable mortgage-related outcomes between eligible and ineligible groups are consistent across regions, both for homes purchased before the Great Recession and for homes refinanced in the early recovery period of 2008 to 2011.

1.2.5 Effects of Internet Essentials on Refinancing

Refinance Originations and Interest Rates. I first study the effect of Internet Essentials on the number of refinance applications and originations. Specifically, I estimate the following equation:

$$\begin{aligned} y_{i,c,t} = {} & \alpha + \beta(Eligible_{i,c,t} \times Comcast_c \times Post_t) + X'_{i,c,t}\Phi \\ & + \rho_1(\lambda_t \times \gamma_c) + \rho_2(Eligible_{i,c,t} \times \lambda_t) + \rho_3(Eligible_{i,c,t} \times \gamma_c) + \epsilon_{i,c,t}, \end{aligned} \tag{1.2}$$

where $y_{i,c,t}$ is the number of refinance originations made by households in eligibility group i in census tract c in year t. I also replace the dependent variable with the number of refinance applications submitted and denial rates to tease out refinancing demand and credit standards, respectively. $Eligible_{i,c,t}$ is a binary indicator for group i's Internet Essentials program eligibility, $Comcast_c$ is a continuous measure of Comcast coverage rates in census tract c, and $Post_t$ indicates years after the introduction of Internet Essentials in 2012. $X_{i,c,t}$ is a vector of eligibility group by census tract by year covariates, which include proxies for house price and income. Census tract by year fixed effects $(\lambda_t \times \gamma_c)$ absorb all census tract-specific trends that are invariant to Internet Essentials eligibility. Similarly, the interaction $Eligible_{i,c,t} \times \lambda_t$ controls for aggregate time-varying differences between eligible and ineligible groups and $Eligible_{i,c,t} \times \gamma_c$ controls for permanent differences between eligible and ineligible groups in each census tract. The parameter of interest, β, captures the remaining variation in $y_{i,c,t}$ which only involves time-varying, within-census tract differences between eligible and ineligible groups. The identifying assumption under this setting, therefore, is that within-census tract differences in refinancing activity between the two groups in high and low Comcast coverage would have trended the same in the absence of Internet Essentials.

In an additional test, I analyze whether households with Internet Essentials are better able to shop around for refinance mortgages and obtain lower interest rates. I replace $y_{i,c,t}$ in equation (1.2) with loan-level interest rates for originated refinance loans between 2008 and 2015. $X'_{i,c,t}$ now includes loan-level covariates such as income, loan amount, race, sex, number of applicants, combined LTV, DTI, credit score, and loan term. The structure of fixed effects are the same as in equation (1.2), and the data comprises a subset of the HMDA source that can be matched to GSE performance filings.

For specifications that involve a count measure as the dependent variable, I use Poisson pseudo maximum likelihood (PPML) regressions to model the data (Gourieroux et al., 1984; Silva and Tenreyro, 2006; Correia et al., 2019). Standard errors are conservatively clustered at the PUMA level to address the possibility that Internet Essentials may have been rolled out in geographic units larger than individual census tracts (e.g., school districts, neighborhoods). All analyses cover the time period between 2009 and 2015 as other major ISPs and government initiatives introduced similar broadband subsidy programs in 2016. Moreover, the federal funds rate started to rise from the zero lower bound at the end of 2015, which would have reduced refinance incentives for marginal households.

Housing Costs. An important testable prediction of refinancing is that housing-related costs should decrease following a refinance. However, it is difficult to directly measure changes in payment burdens at the household level as an old mortgage cannot be linked to the refinanced mortgage using HMDA data. In this section, I use annual survey responses from the ACS to quantify Internet Essentials' effect on housing costs for both homeowners and renters. I estimate the following equation using survey responses geographically identified at the PUMA level:

$$\begin{aligned} m_{i,p,t} = \alpha &+ \beta(Eligible_{i,p,t} \times Comcast_p \times Post_t) + Z'_{i,p,t}\Phi \\ &+ \rho_1(\lambda_t \times \gamma_p) + \rho_2(Eligible_{i,p,t} \times \lambda_t) + \rho_3(Eligible_{i,p,t} \times \gamma_p) + \epsilon_{i,p,t}, \end{aligned} \tag{1.3}$$

where $m_{i,p,t}$ is either the natural logarithm of monthly mortgage payments (rent payments) or the mortgage to income ratio (rent-to-income ratio) for household i in PUMA p in year t. $Eligible_{i,p,t}$ is an eligibility indicator that now varies for each household i following the definition outlined in 1.2.3. In an alternative specification, I restrict the ineligible group further to households with income below 170 percent of the FPL but without a school-aged child. This step further aligns the treatment and control groups in terms of observable characteristics while maintaining variation in program eligibility. $Comcast_p$ indicates whether more than 90 percent of PUMA p's population is covered by Comcast (control group with less than 10 percent in coverage). The redefinition of $Comcast_p$ is necessary because PUMAs are on average 10 times larger than census tracts in terms of population; PUMAs with medium levels of coverage may confound the results as they might be areas with more than one major ISP in operation (including Comcast).[17] $Z_{i,p,t}$ is a vector of household-specific covariates obtained from relevant sections of the ACS. To mitigate the effect of new homeowners that may have obtained their first mortgage at lower rates, I restrict the sample to households that moved into their current residence more than three years prior to the response period. Lastly, I relax the urbanicity requirement in order to reduce the increased demand on the data arising from PUMA level variation. Concerns of confounding trends as a result of this adjustment are low due to the cleaner identification of household-level eligibility status. Multi-way fixed effects absorb any variation that might threaten the validity of the identification strategy. Standard errors are clustered at the PUMA level.

[17] In unreported results, I verify that the regression results do not change materially when using the continuous measure of Comcast coverage as in equation (1.2).

1.3 Results

1.3.1 Main Results

Refinance Outcomes. I first estimate the effect of Internet Essentials on refinance outcomes (applications, originations, and denial rates) at the eligibility group level. Column 1 in Table 1.6 presents triple differences estimates on refinance originations. I find that the availability of Internet Essentials increased the number of new mortgages originated to eligible households by 6 percent per year, relative to an average of 6 mortgages. These results are statistically significant at the 5 percent level. Figure 1.8 graphically illustrates these results by plotting time-varying triple difference estimates of treatment effects. I find no evidence of non-parallel trends in the pre-treatment period, confirming the validity of a granular identification strategy that exploits variation between groups and across census tracts. Importantly, the coefficient estimates on refinance originations steadily grow over the early years of the program and become statistically significant in 2013 and 2014. The gradually increasing trend also mirrors the subscriber growth pattern between 2012 and 2015 (Comcast Corporation, 2016). The treatment effect falls marginally and becomes insignificant in 2015, corresponding to the eventual slowdown in aggregate refinancing demand.

I also do not find evidence that the increase in refinance originations is associated with suboptimal application behavior. As low-income households are more likely to have creditworthiness that is marginally sufficient to qualify for a mortgage, it is possible that the growth in refinance originations masks an increase in costly denials. I indirectly test the hypothesis that access to the internet can have the unintended consequence of disseminating misinformation or inflating the perceived likelihood of approval using applications and denial rates data from HMDA. In column 2 of Table 1.6, I show that the number of applications also increases by 6 percent and that the coefficient is statistically significant at the 1 percent level. Column 3 corroborates these results that there is no effect on refinance denial rates relative to a pre-treatment average of 31 and 41 percent for eligible and ineligible groups, respectively. These results imply that internet access does not induce suboptimal refinancing behavior. Moreover, banks and mortgage lenders do not seem to adjust lending standards in response to the increase in applications, which is plausible given the comparison of outcomes for two similar groups within a census tract.

Internet Essentials also did not induce borrowers to obtain more favorable interest rates conditional on approval. Column 4 shows a non-significant effect of treatment on interest rates controlling for a rich set of loan-level covariates. This can be explained by the relative uniformity of conventional mortgages compared to other types of programs. In addition, fintech lenders did not have a large market share in retail mortgages during this period, which may have led to higher frictions for online rate-shopping activities (Figure 1.9). Even if online rate search tools are utilized by homeowners, online lenders tend to charge similar or higher interest rates than their brick-and-mortar counterparts to compensate for improved convenience (Buchak et al., 2018).

The monetary savings from refinancing are economically substantial, especially for low-income households that have most of their wealth tied to home equity. The average homeowner in my sample that purchased a home between 2004 and 2008 had a mortgage principal of around \$120,000 and an interest rate of 6.2 percent at the time of origination. Applying the prevailing interest rate of 4 percent for comparable loans between 2012 and 2015, each household that refinanced its mortgage would have saved \$110 dollars a month before any adjustments. These households still come out ahead by around \$100 after accounting for the cost of Internet Essentials, which corresponds to about 5 percent of disposable income for the average household in this group. More importantly, the lifetime savings for an average refinance loan can be up to \$29,000, or \$18,000 after discounting over time and adjusting for possible closing costs.[18] These lifetime savings account for around one-third of the median net worth of all households and about 10 percent of the net worth of low-income homeowners residing in owner-occupied units (Survey of Consumer Finances, 2013; Wolff, 2016).

I also estimate the aggregate economic impact of Internet Essentials to be large and persistent. A 6 percent increase in the number of refinance originations, off a base of 13,000 annual originations for the treated group prior to 2012, corresponds to 780 additional refinances per year (total origination volume of $100 million per year). Based on the aforementioned conservative measure of household wealth gains ($18,000), Internet Essentials generated $55 million in aggregate household savings through refinancing between 2012 and 2015. These results importantly ignore the effect on non-urban households, and the the upper bound of national savings attributable to Internet Essentials is around $100 million.[19] Taking stock, these aggregate savings almost directly offset the $110 million that Comcast invested into public service announcements to advertise the program during this period. Even if we assume that Comcast breaks even on each subsidized line, the mortgage cost savings combined with other documented economic benefits such as increased employment outcomes (Zuo, 2021) imply that providing subsidized broadband can indeed be a desirable government policy.

Mortgage Payments. I further test whether Internet Essentials indeed led to lower mortgage payments. This is an important empirical exercise given the incidence of suboptimal refinancing behavior particularly among low-income households (Agar-

[18] To calculate the present value, I use a discount rate of 4 percent and adjust the savings downward by an additional 15 percent to account for marginal taxes, closing costs and the probability of moving. Note that closing costs can often be waived for low-income borrowers through federal and state grant programs. Using a more conservative set of parameters from Agarwal et al. (2013) and Keys et al. (2016) would further reduce the estimated savings to $15,000, which is still very high for this group of homeowners.

[19] Urban census tracts account for around 54 percent of Comcast's coverage area by population. Assuming that the treatment effect of the program would have been the same (or half as effective) in non-urban census tracts, the upper (lower) bound of mortgage payment savings is $100 million ($78 million).

wal et al., 2016). Even if mortgage payments decrease, the true effect of actual savings may be lower than the 14 percent derived from average interest rate differences due to origination costs, taxes, or fluctuations in appraisal value. Table 1.7 shows the results from estimating equation (1.3). Panel A uses a control group of all eligibles (higher income, no school-aged child, or both). I find that Internet Essentials decreased mortgage payments in treated areas by 2.5 percent and the mortgage to income ratio by 1.5 percent. The results are statistically significant and are robust to the inclusion of control variables for demographics (e.g., age, race, gender, educational attainment) and economic characteristics (income, home value). Additionally, Panel B improves on the identification by comparing mortgage payment outcomes between low-income households with at least one school-aged child and low-income households without a school-aged child. This specification yields similar coefficients for mortgage to income ratio and an even larger effect on mortgage payments of 3.8 percent. In Figure 1.10, I verify that the point estimates on log mortgage payments, the cost measure of choice, are not statistically significant prior to 2012. The point estimates generally decrease over time after Internet Essentials is introduced and becomes statistically significant in 2014 for panel (a). However, I do not find a statistically significant effect in any other year under either specification. This fact, in conjunction with the negative and statistically significant effect on the baseline triple-differences analysis, can be explained by the relative infrequency of refinancing events among low-income homeowners and data limitations.

The magnitude of treatment effects in Table 1.7 provide important baseline estimates for the monetary savings from refinancing. The average pre-treatment mortgage payment for treated households is around \$700, which is consistent with the statistics obtained from HMDA. A 4 percent decrease in payment corresponds to \$30 in monthly savings or \$5,500 in adjusted present value terms. This serves as the lower bound for the treatment effect of Internet Essentials on mortgage payments, as the

ACS does not directly collect information about mortgage refinancing activity. Even if we take the estimates at face value, I argue that $30 a month could make a large difference in financial health when accumulated over several decades. This is because disposable income and discretionary savings for low-income households are extremely low. In fact, 32.8 percent of households that are income-eligible for Internet Essentials reported to be "food insecure," which means they did not have access to enough food for an active, healthy life for all household members (Coleman-Jensen et al., 2016).

1.3.2 Mechanisms

In this section, I analyze the mechanisms through which expanding broadband access improves refinancing outcomes for low-income households. Internet Essentials' unique empirical setting provides testable predictions for whether the positive effect of broadband on refinancing is a result of the rise in online lending or improved access to traditional mortgage services. Moreover, I study two competing explanations for higher refinancing demand — the income effect of broadband connectivity and reduced informational frictions.

Lending Channels and Financial Inclusion. Access to traditional financial services such as a mortgage is particularly challenging for the 20 percent of American households that are classified as underbanked (Federal Deposit Insurance Corporation, 2013).[20] Between 2008 and 2016, more than 6 percent of bank branches closed throughout the nation, making it difficult for households living in underserved areas to refinance their mortgages conveniently (National Community Reinvestment Coalition). Branch closure rates are in fact more pronounced in urban areas with relatively high internet access levels (Jiang et al., 2022); for instance, Comcast cities such as

[20] A household is underbanked if it used alternative financial services (money orders, check cashing, remittances, payday loans, refund anticipation loans, rent-to-own services, pawn shops loans, or auto title loans) from non-bank providers in the preceding 12 months.

Chicago (13 percent), Philadelphia (18 percent), and Detroit (16 percent), as well as no Comcast cities such as New York (11 percent), Dallas (8 percent), and Las Vegas (17 percent) experienced significant branch closures during this period.

I outline two supply-related predictions for the refinancing activity of households with limited access to bank branches. First, broadband access may encourage homeowners to refinance through online (fintech) lenders that are more efficient in processing applications. Alternatively, broadband access can facilitate the refinance process via traditional banking relationships by reducing shadow costs or search frictions.

Table 1.8 empirically tests these two hypotheses. In column 1, I replace $y_{i,c,t}$ in equation (1.2) with the fraction of fintech originations to all refinance originations. I do not find any effect of Internet Essentials on fintech relationships, which can partially be explained by the relatively low levels of fintech penetration across eligible and ineligible income groups during this period (4.3 percent and 7.2 percent). This fact is also supported by Figure 1.9, which shows that Google search activity for the top fintech lenders remained muted until 2015.

In columns 2 to 4, I estimate equation (1.2) for refinance originations after dividing the sample of census tracts into three groups based on the number of physical bank branches within 2 miles of the population center. I find that the treatment effect is largest (9.1 percent) when comparing the refinancing gap across Comcast and no Comcast census tracts in the bottom quintile of bank branch density (average of 4.12 branches around population center). The treatment effect is smaller and not significant for the middle quintile, and importantly, I find no effect when comparing census tracts with the highest levels of branch access. These results imply that Internet Essentials had the largest impact in areas where households face high shadow costs of refinancing. This is consistent with the findings of Argyle et al. (2020) that low bank branch access is associated with higher search costs and worse consumer financial outcomes. Furthermore, low-income households are likely to be constrained in their ability to make long-distance branch visits as they are generally in service, natural resources, maintenance, and construction occupations that exhibit limited flexibility in work schedules.[21] Thus, I demonstrate that broadband improves low-income households' refinancing outcomes by reducing the shadow costs of accessing traditional brick-and-mortar lenders, which are still considered the main source of credit for disadvantaged populations.

Determinants of Refinancing Demand. In this section, I disentangle the possible determinants of increased refinancing demand following Internet Essentials. Zuo (2021) shows that the program led to increased employment and income for eligible households residing in Comcast areas. Given this, improved financial health may have enabled refinancing for households that previously did not have enough savings or work flexibility to cover monetary origination costs as well as shadow costs. For this hypothesis to hold true, it must be the case that income for eligible homeowners, which account for less than half of all households in this income group, indeed increased as a result of the program. I test whether the results on income from Zuo (2021) hold when restricting the sample to homeowners only. Specifically, I replace the dependent variable in equation (1.3) with the log of income conditional on having a mortgage and being employed. This is because refinancing is only relevant for employed households in most circumstances. I use the preferred specification that assigns the control group as low-income households that are ineligible due to the absence of a school-aged child. Column 1 of Table 1.9 shows that Internet Essentials in fact did not have any effect on income for employed households with a mortgage. This result rules out the possibility that refinancing demand increased due to the reduction of opportunity costs.

An alternative explanation for increased refinancing demand is that Internet Essentials bridged the large gap in digital and financial literacy between connected and

[21] Bureau of Labor Statistics (2014).

unconnected households. In particular, survey results indicate that online access of banking and financial services is much higher among households with high digital skills (60 percent) than households with low digital skills (39 percent) (Horrigan, 2019). Recognizing the importance of training programs that help households transition daily activities online, Comcast invested $300 million into digital literacy initiatives that were accessed by 30 percent of subscribers.(Comcast Corporation, 2016). The programs, which were offered free of charge through multiple outlets, covered a wide range of topics on digital readiness (e.g., internet security and e-mail) as well as general well-being (e.g., employment, social services, and personal finance).

I test whether the refinancing growth among treated households can be explained by an increase in digital and financial literacy (measured by educational attainment). Columns 2 to 4 of Table 1.9 estimate regression (1.2) with refinance origination counts as the dependent variable. I again divide the census tracts into three groups based on the fraction of the population with a high school degree or higher. Column 2, which compares the refinancing gap between eligible and ineligible households in urban census tracts with low levels of educational attainment as of 2011, reveals a positive and statistically significant coefficient of 12.5 percent. I find no treatment effect in the middle group of census tracts, and a positive and significant coefficient of 6 percent for high literacy census tracts.

In order to tease out these channels more directly, I also estimate regression (1.3) using log mortgage payment as the dependent variable and then dividing the sample of ACS respondents into low (less than high school degree) and high (at least high school degree) digital and financial literacy groups. Instead of focusing on geography-specific education levels, I focus on household-level variation in educational attainment in this specification. Columns 5 to 8 provide further support of this channel: Internet Essentials reduced mortgage payments by 5.4 to 8.3 percent among low literacy groups, but had no effect on high literacy households.

Taken together, these results confirm that Internet Essentials increased refinancing demand by improving the digital and financial literacy of low-income households. Households with higher ex-ante levels of digital and financial literacy were not differentially impacted by Internet Essentials, implying that if desired, they would have refinanced one way or another even without an at-home broadband connection.

1.4 Robustness and Falsification Tests

Alternative Measure of Refinancing. While loan counts provide the most direct and comprehensive measure of refinancing activity, it importantly cannot shed light on how refinancing inequality evolves relative to a stock of existing, current mortgages. To address this, I analyze the evolution of prepayment behavior for home purchase mortgages originated between 2004 and 2008 in a two-period model. I estimate the following equation:

$$\begin{aligned} prepay_{i,c,t} = \alpha &+ \beta(Eligible_{i,c,t} \times Comcast_c \times Post_t) + Y'_{i,c,t}\Phi \\ &+ \rho_1(\lambda_t \times \gamma_c) + \rho_2(Eligible_{i,c,t} \times \lambda_t) + \rho_3(Eligible_{i,c,t} \times \gamma_c) + \epsilon_{i,c,t}, \end{aligned} \tag{1.4}$$

where $prepay_{i,c,t}$ is a binary indicator for whether loan i in census tract c has prepaid by year $t \in \{2011, 2015\}$. $Eligible_{i,c,t}$ now indicates whether loan i qualifies for Internet Essentials at the time of origination, and I assume that eligibility status stays constant between origination and 2015. To address the concern that households with marginally higher income between 2004 and 2008 may have subsequently qualified for the program by 2012, I construct an additional control group with annual income between 185 percent and 370 percent of a seven-person household ($55,000 to $110,000). $Y'_{i,c,t}$ is now a vector of loan-specific covariates, which includes income, race, sex, number of applicants, interest rate at origination, loan-to-value ratio, debt-to-income ratio, credit score, loan amount, and mortgage tenure. Census tract by year fixed effects ($\lambda_t \times \gamma_c$) absorb all census tract-specific trends that are invariant to Internet Essentials eligibility and the interaction $Eligible_{i,c,t} \times \lambda_t$ controls for aggregate time-varying differences between eligible and ineligible groups. Similarly, $Eligible_{i,c,t} \times \gamma_c$ controls for permanent differences between eligible and ineligible groups in each census tract.

Column 1 of Table 1.10 shows that the prepayment probability of a conventional mortgage originated by a low-income household before the Great Recession increased by 3.3 percent as a result of Internet Essentials. The effect is economically large given the average pre-treatment prepayment propensity of 42 percent, and is statistically significant at the 1 percent level.

Direct measurement of prepayment outcomes also allows me to compute how much of the reduction in refinancing inequality between the top and bottom income deciles between 2011 and 2015 can be attributed to Internet Essentials. First, the effect of prepayment activity estimated in Table 1.10 implies that Internet Essentials can explain up to 10 percent of the growth in prepayment for the lowest income decile.[22] In addition, back of the envelope calculations suggest that the program reduced the gap in refinancing activity between the top and bottom income deciles by 14 percent. These estimates reflect an upper bound as the reduction in refinancing gap is largely a result of mechanical convergence over time.[23]

Alternative Eligibility Thresholds. Internet Essentials eligibility is importantly based on household income and family composition, the latter of which I

[22] A 3.3 percent increase off a base of 65 percent implies a 2 percentage point increase in prepayment. I then divide this number by the total prepayment growth by this group during this period (23 percent). Note that the slight difference in base prepayment propensities compared to the regression results is due to the use of static income deciles for illustrative purposes.

[23] At the end of 2011, there was a 23 percent gap in the fraction of pre-crisis mortgages refinanced between the bottom and top income deciles in urban census tracts with high Comcast coverage (65 vs. 88 percent). The same gap was reduced to 9 percent (89 vs. 98 percent) by the end of 2015. I take the ratio of the aforementioned prepayment growth and the reduction of the gap (14 percent).

cannot directly measure from the HMDA or GSE Data. While the fact that income conditional on homeownership does not increase helps rule out the possibility of an ineligible household becoming ineligible again, my analysis still suffers from the concern that the choice of income thresholds does not precisely identify truly eligible households. As such, the validity of my findings would be undermined if I find positive treatment effects on refinancing activity when using groups of lower-middle income households that are both unlikely to be impacted by Internet Essentials. In column 2 of Table 1.10, I show that assigning a placebo treated group at the 5-6 person income threshold and control group at the 7-8 person income threshold does not yield statistically significant effects on mortgage refinancing. The disappearance of an effect confirms that my criteria coincides with actual eligibility and that income-based differences in refinancing trends alone cannot explain the findings. In unreported analysis, I also verify that expanding the control group to households within the 5-7 person income threshold does not materially change the results.

Placebo ISPs. Internet Essentials was the only broadband subsidy program of its kind until 2016. After that, other major ISPs as well as federal and state governments introduced similar initiatives to bridge the digital divide. These multilateral efforts were made more prominent and permanent following the COVID-19 pandemic. Because of the uniqueness of Internet Essentials between 2012 and 2015, the causal estimates on refinancing should disappear when I assign AT&T or Charter as placebo program providers. To tes this, I compute coverage rates $AT\&T_c$ and $Charter_c$ at the census tract level and re-estimate equation 1.2. Columns 4 and 5 of Table 1.10 report the results. Indeed, instituting a placebo broadband program in high AT&T and high Charter areas do not yield any effect on refinance originations.

Rental Costs. Rentals are the prominent alternative housing tenure choice for households. Unlike mortgages, rent payments are typically contractual and regulated by local housing authorities. Renting also does not allow households to build wealth through home equity, meaning that long-term gains from converting into lower rent payments are less likely to be consequential for low-income households. As such, outcomes on rental payments should not change as a result of Internet Essentials. I test this prediction using ACS data on renters and confirm that households do not take advantage of Internet Essentials to reduce their rent payments (Table 1.10, column 6).

Census Tract Characteristics. Despite reports of Internet Essentials' rapid growth nationwide, I cannot directly observe program take-up at the household, loan, or geographic area level.[24] As an additional falsification test, I analyze whether the treatment effects on refinancing activity are concentrated among census tracts that are likely to have a large pool of new program subscribers. First, in columns 2 to 4 of Table 1.11, I show that census tracts with a higher fraction of owner-occupied households with children between ages 6 and 18 — one of the main criteria for program eligibility — contribute to the entirety of treatment effects. Second, it is plausible that refinancing demand is correlated with housing cost burdens as the impact of payment savings are largest. I measure census tract-level housing cost burdens as the fraction of homeowners who pay more than 30 percent of income on mortgages as of 2011. Since this measure is calculated regardless of income, it also partially captures differences in local house prices. Columns 5 to 7 report the results: only the treatment effect of refinance originations at the top quartile (16.3 percent) is statistically significant. Lastly, I test for heterogeneous effects across census tracts within PUMAs with varying levels of broadband subscription rates as of 2013, the first year this question was asked in the ACS. Again, a statistically significant treatment effect of 9.2 percent is only present when comparing census tracts in the top quartile of broadband subscription rates. This result provides suggestive evidence that areas with

[24] Zuo (2021) estimates a program take-up rate of 10.6 percent across all Comcast areas between 2012 and 2015.

resilient existing broadband infrastructure (such as stronger advertising campaigns, better equipment efficiency, or resilient social networks) benefited the most from the program. Conversely, the finding also implies a relative inefficiency in less connected areas that can be addressed through increased targeting efforts.

1.5 Conclusion

Failing to refinance a mortgage when it becomes profitable to do so leads to large welfare losses. This phenomenon is particularly prominent among low-income households and has contributed to the growing wealth inequality in recent decades. In this paper, I study whether disparities in access to the internet explains suboptimal refinancing behavior by exploiting a natural experiment that brought broadband to more than low-income 750,000 households between 2012 and 2015. Using an identification strategy that accounts for geographic, temporal, and household-level variation in program availability, I find a strong and positive effect on refinancing outcomes that lead to a decrease in mortgage cost burdens. The economic significance of the results are large and persistent, resulting in total savings that correspond to 10 percent of net worth of low-income households. The effects are driven by areas that are underbanked as well as areas with low levels of digital and financial literacy. I conduct various robustness and falsification tests to confirm that my findings are indeed driven by increased access to broadband internet.

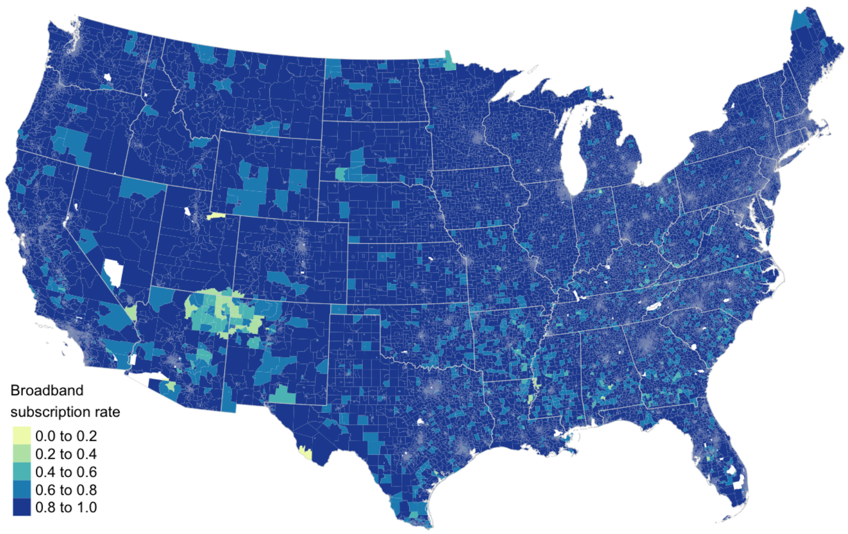
This paper provides important implications for monetary policy, mortgage contract design, and infrastructure policy. First, the pass-through of accommodative monetary policy via refinancing may be hindered by shadow costs that differentially affect households with or without internet access. Since the digital divide persists along the income dimension and in less developed areas, the consequences of failing to refinance for disadvantaged groups will be amplified during economic downturns.

Moreover, a housing market that is dominated by fixed-rate mortgages exacerbates wealth inequality by placing the burden of refinancing solely on households. Over the past several decades, low-income and minority families have been stymied by the mismatch between following the path to homeownership and the lack of ability or resources to refinance when it becomes optimal to do so. To address this, the government and financial institutions should consider developing alternative mortgage products that target these populations and dynamically induce refinancing behavior. Lastly, large-scale efforts to get Americans connected to broadband should continue via improvements in affordability and expanded physical access.

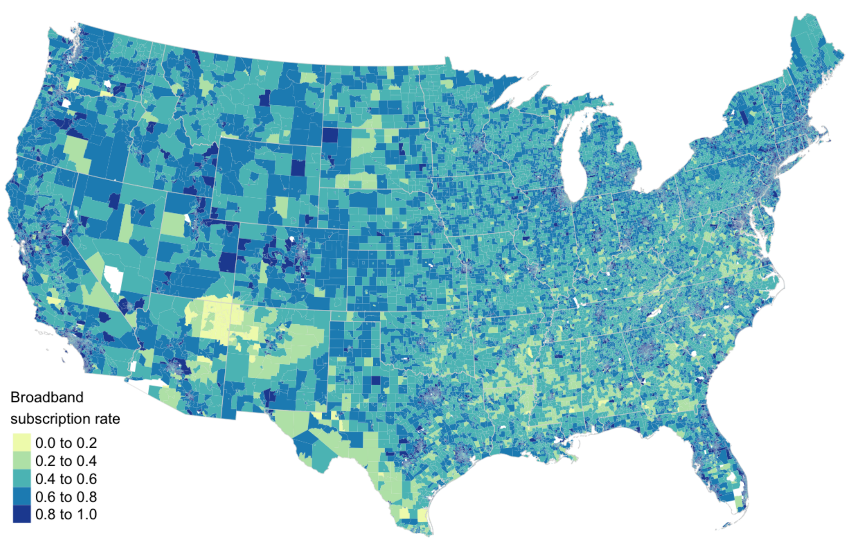
Access to high-speed internet is one of the most prominent equalizing forces in the modern era. As technology continues to evolve, the new front of financial inclusion will depend less on introducing branches and ATMs to neighborhoods but more on connecting people via devices and applications. While this paper addresses the internet's key role bridging the wealth gap in the context of mortgages, additional consideration should also be given to other aspects of household finance such as savings and investment behavior.

APPENDICES

1.A Figures and Tables



(a) Household income above $75,000



(b) Household income below $35,000

Figure 1.1: Broadband Access in the United States

Note: This figure plots the fraction of high- and low-income households with a broadband internet subscription at the census tract level. Annual household income is in 2019 inflation-adjusted dollars. Source: 2019 ACS 5-year estimates.

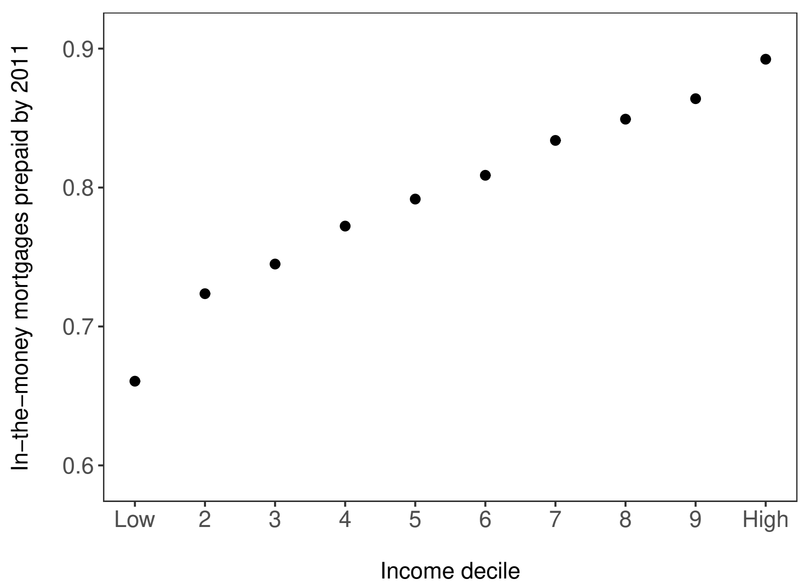


Figure 1.2: Household Income and Refinancing Inequality

Note: This figure plots the relationship between household income and mortgage prepayment. For each income decile of households that originated a conventional mortgage sold to Fannie Mae or Freddie Mac between 2004 and 2008, I calculate the total volume of mortgages with above-median interest rates and credit quality metrics (combined LTV, DTI, and credit score). Then, I compute the fraction of these mortgages that were voluntarily prepaid (by volume) on or before 2011. The sample consists of loans in urban central metro areas.
Source: HMDA, Fannie Mae and Freddie Mac loan performance files, and author's calculations.

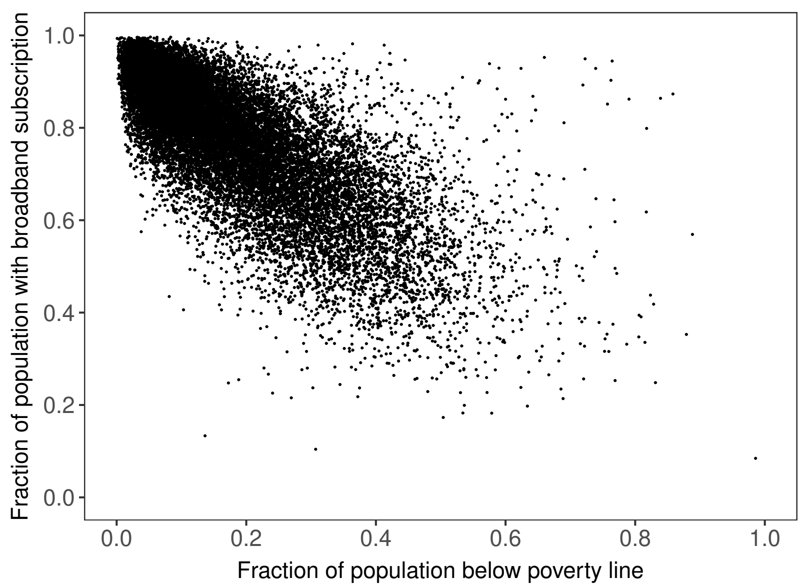


Figure 1.3: The Digital Divide in Large Central Metro Counties

Note: This figure shows broadband inequality in large central metro counties with high levels of ISP coverage. The x-axis is the fraction of a census tract's population living below the poverty line, and the y-axis is the fraction of the population with a high-speed broadband subscription at home. Source: NCHS, 2017 ACS 5-year estimates.

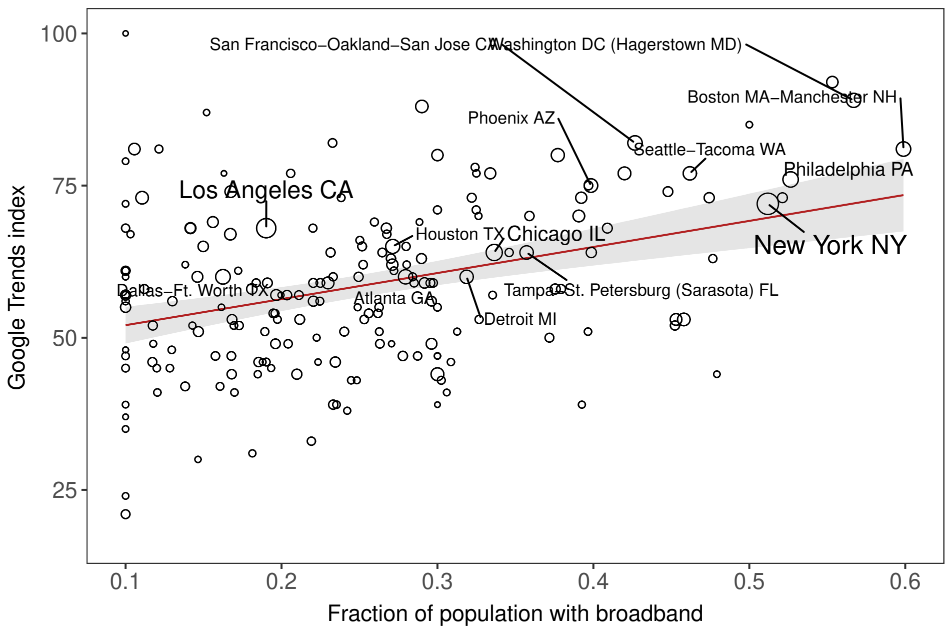


Figure 1.4: Broadband Access and Refinancing Demand

Note: This figure plots the relationship between broadband connectivity and refinancing demand. Google Trends search data for relevant keywords ("refinance," "refinance rates," "mortgage refinance," and "mortgage rates") are compiled for each metropolitan area between 2012 and 2015. Broadband subscription data (at least 10 Mbps download speed) is compiled at the county level as of December 2011. I match these two data sources and calculate a weighted broadband index at the metropolitan area level. The shaded region represents 95 percent confidence intervals for the linear fitted line. The size of each observation indicates the size of each area, and locations with more than 2 million housing units are labeled.
Source: Google, FCC Form 477, geography crosswalk file from Jacob Schneider.

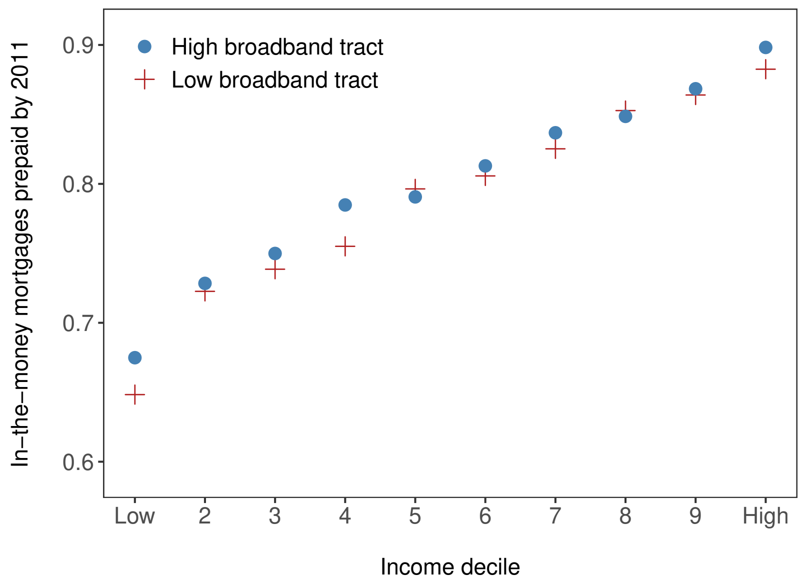


Figure 1.5: Refinancing Inequality and Broadband Access

Note: This figure separately plots the relationship between household income and mortgage prepayment in high- and low-broadband census tracts. "High broadband tract" and "low broadband tract" are defined as census tracts that had below 40 percent and above 60 percent coverage of broadband subscription rates as of December 2011, respectively. Income deciles are constructed using conventional mortgages originated and sold to Fannie Mae and Freddie Mac between 2004 and 2008. I restrict the sample to mortgages with above-median interest rates and credit quality metrics (combined LTV, DTI, and credit score) at time of origination. I plot the fraction of these mortgages that were voluntarily prepaid (by volume) on or before 2011. The sample consists of loans urban central metro areas.
Source: HMDA, Fannie Mae and Freddie Mac loan performance files, FCC Form 477, and author's calculations.

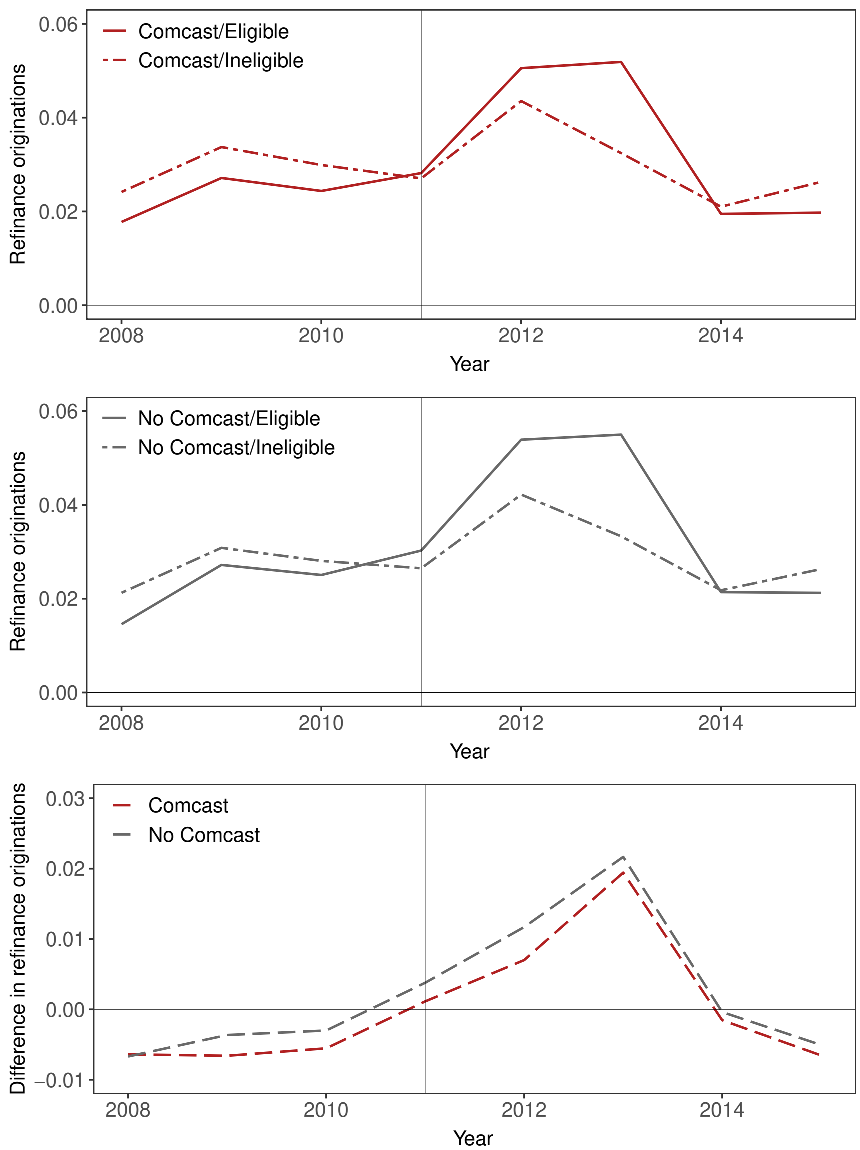
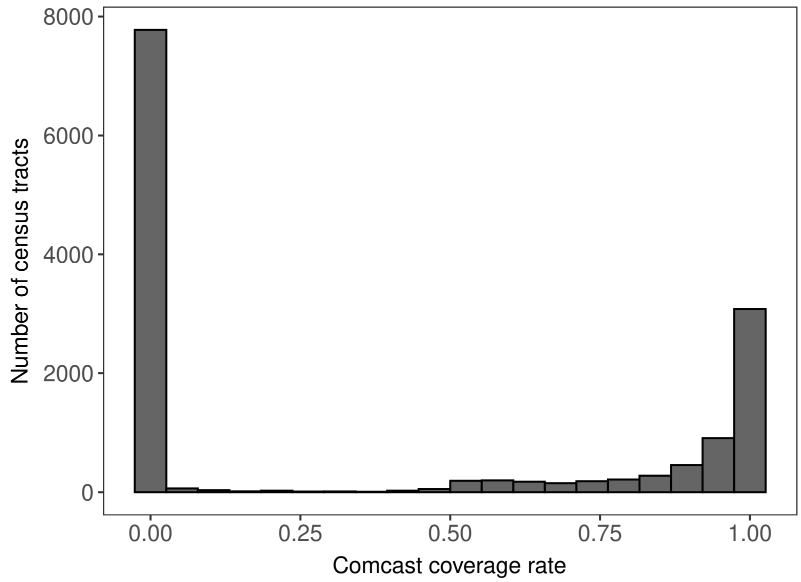
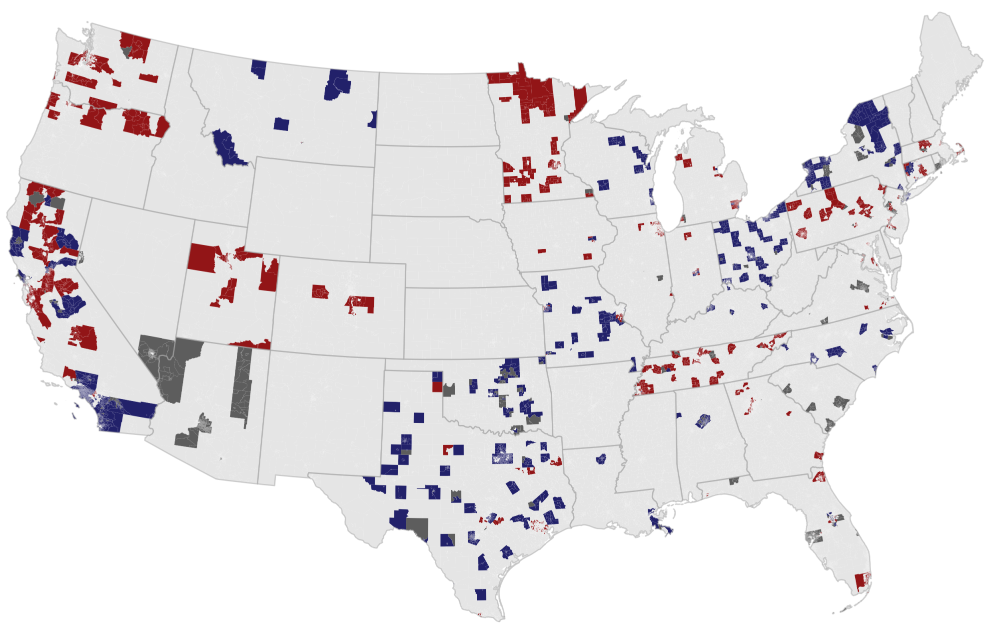


Figure 1.6: Unconditional Trends in Refinancing Activity

Note: This figure illustrates the triple differences empirical design by plotting the unconditional refinancing trends between eligibles and ineligibles across Comcast and no Comcast census tracts. Refinance originations is measured as the number of loans originated by eligibility group divided by the imputed stock of owner-occupied households with a mortgage, and is residualized with respect to proxies for house prices (value of newly originated mortgages) and economic conditions (income). The bottom panel plots the difference in the two series by Comcast and no Comcast status. The sample covers large central metro census tracts.
Source: HMDA, 2011 ACS 5-year estimates, 2010 Decennial Census.



(a) Histogram of Comcast Coverage Rates



(b) Map of Comcast and No Comcast Census Tracts

Figure 1.7: Comcast Coverage Rates

Note: This figure plots the statistical and geographical distributions of Comcast coverage in large central metro census tracts. For each census tract, I first calculate the fraction of population with Comcast access. The final coverage rate takes the average of coverage rates in December 2011 and December 2014. THe top panel shows the distribution of comcast coverage rates. The bottom panel illustrates Comcast (red), no Comcast with AT&T and Charter (blue), and other no Comcast census tracts (dark grey).
Source: NTIA SBI, NCHS.

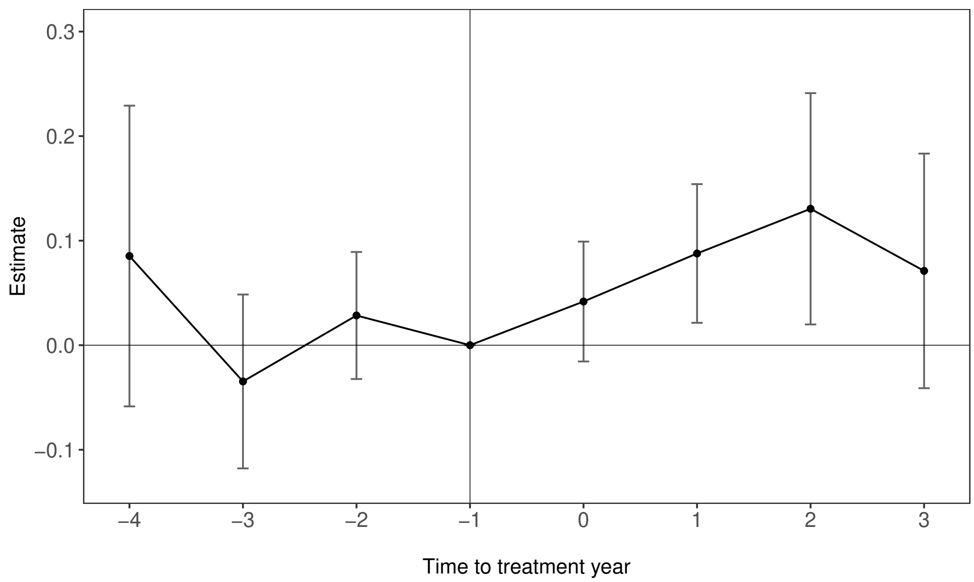


Figure 1.8: Event Study Estimates for Refinance Originations

Note: This figure plots dynamic triple difference estimates (β_t) and 95 percent confidence intervals for the number of refinance originations. The estimating equation is:

$$
\begin{aligned}
y_{i,c,t} =& \alpha + \sum_t \beta_t(Eligible_{i,c,t} \times Comcast_c \times Year_t) + X'_{i,c,t}\Phi + \rho_1(\lambda_t \times \gamma_c) + \rho_2(Eligible_{i,c,t} \times \lambda_t) \\
&+ \rho_3(Eligible_{i,c,t} \times \gamma_c) + \epsilon_{i,c,t}, \quad t \in \{-4, -3, -2, 0, 1, 2, 3\}.
\end{aligned}
$$

The sample spans the period between 2008 and 2015. The interaction term in the final pre-treatment period (2011) is omitted. Robust standard errors are clustered at the PUMA level.
Source: HMDA, ACS IPUMS microdata.

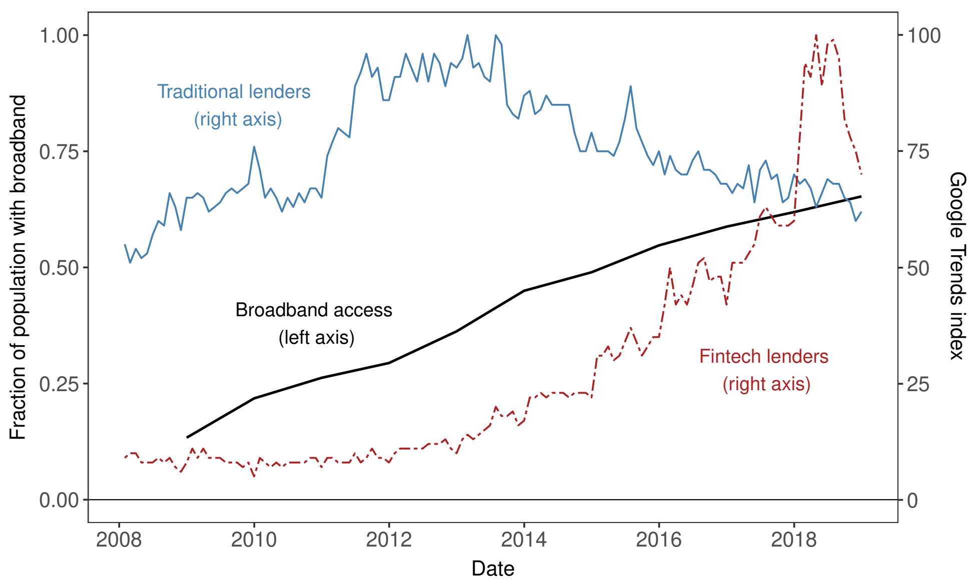
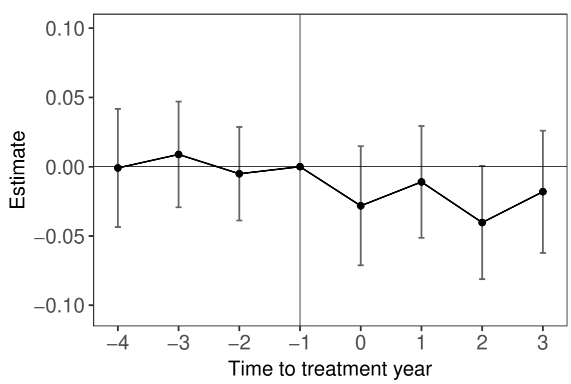
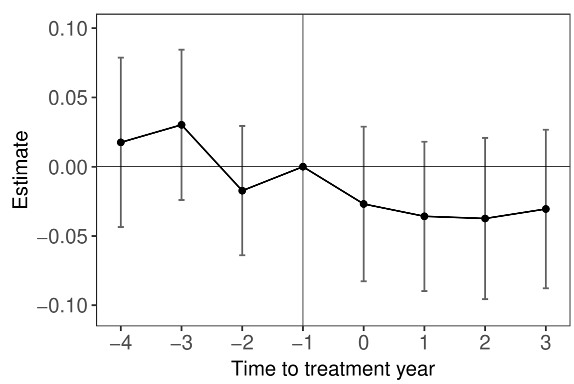


Figure 1.9: Trends in Online Search for Refinancing

Note: This figure plots the evolution of online search trends for traditional and fintech mortgage lenders. Google Trends search data for the top 10 traditional lenders and top 10 fintech lenders by origination volume are plotted each month from 2008 to 2018. The search indices are normalized relative to a maximum of 100 during the study period. Fintech lender classification follows Buchak et al. (2018) and Fuster et al. (2019). National broadband subscription data are computed using county level annual subscription estimates and housing unit counts. Broadband is defined as wireline connections with a minimum download speed of 10 Mbps.
Source: Google, FCC Form 477.



(a) All ineligibles



(b) Low-income ineligibles

Figure 1.10: Event Study Estimates for Mortgage Costs

Note: This figure plots dynamic triple difference estimates (β_t) and 95 percent confidence intervals for log mortgage payment. Panel (a) includes all ineligibles as the control group, while panel (b) focuses on low-income ineligibles as the control group. The estimating equation is:

$$\begin{aligned} m_{i,p,t} =& \alpha + \sum_t \beta_t (Eligible_{i,p,t} \times Comcast_p \times Year_t) + Z'_{i,p,t}\Phi + \rho_1(\lambda_t \times \gamma_p) + \rho_2(Eligible_{i,p,t} \times \lambda_t) \\ &+ \rho_3(Eligible_{i,p,t} \times \gamma_p) + \epsilon_{i,p,t}, \quad t \in \{-4, -3, -2, 0, 1, 2, 3\}. \end{aligned}$$

The sample spans the period between 2008 and 2015. The interaction term in the final pre-treatment period (2011) is omitted. Robust standard errors are clustered at the PUMA level.
Source: HMDA, ACS IPUMS microdata.

Table 1.1:

Internet Essentials and Home Internet Use

This table provides summary statistics on demographic characteristics of Internet Essentials subscribers and information on internet usage. The data are collected from anonymous surveys administered by the Comcast Technology Research & Development Fund between 2012 and 2014. (Comcast Corporation, 2016; Horrigan, 2014). All estimates are based on survey respondents and may not necessarily represent the head of household.

	Estimate
Subscriber household characteristics	
Average age	39
Average household size	4
Female (%)	74
Married (%)	46
High school diploma or less (%)	51
Income less than $40,000 (%)	78
Race/ethnicity	
White (%)	44
Hispanic (%)	43
Black or African-American (%)	33
Demand factors and usage	
Children's schoolwork (%)	98
Finding general information (%)	92
E-mail (%)	80
Social networking (%)	71
Paying bills (%)	63
Access to banks and financial institutions (%)	65
Access to government services (%)	52
Access to employment/job search (%)	49

Table 1.2:

Income Thresholds for Internet Essentials Eligibility

This table reports changes in annual income thresholds for Internet Essentials eligibility, which is in turn determined by household size and poverty status. I define households with annual income less than 185 percent of the FPL for a three-person household as eligible. For the ineligible group, I assign the minimum and maximum income as 185 percent of the FPL for a five-person and six-person household, respectively. All thresholds are shown in dollars (thousands).

	Eligible		Ineligible	
Year	Min	Max	Min	Max
2008	0	32.56	45.88	52.54
2009	0	33.87	47.71	54.63
2010	0	33.87	47.71	54.63
2011	0	34.28	48.42	55.48
2012	0	35.32	49.97	57.30
2013	0	36.13	51.01	58.44
2014	0	36.61	51.63	59.15
2015	0	37.17	52.56	60.26
Average	0	34.98	49.36	56.55

Table 1.3:

Urban Metropolitan Statistical Areas by Comcast Coverage

This table lists the top 15 Comcast and no Comcast MSAs by population served. I classify census tracts with more than 50 percent coverage between 2011 and 2014 as Comcast and less than 50 percent coverage as no Comcast. For each MSA, I tally the number of Comcast and no Comcast census tracts and aggregate their respective populations obtained from the 2010 Census. The resulting MSAs are then ranked by population size.

		Census tracts	2010 population (millions)
Comcast			
1	Chicago–Naperville–Joliet, IL	184	1.935
2	Minneapolis–St. Paul-Bloomington, MN-WI	237	1.446
3	San Jose–Sunnyvale–Santa Clara, CA	113	1.340
4	Oakland-Fremont–Hayward, CA	231	1.312
5	Sacramento–Arden-Arcade–Roseville, CA	111	1.193
6	Miami–Miami Beach–Kendall, FL	91	1.182
7	Houston–Sugar Land–Baytown, TX	94	0.994
8	Philadelphia, PA	55	0.941
9	Seattle–Bellevue–Everett, WA	166	0.927
10	Pittsburgh, PA	227	0.918
11	Salt Lake City, UT	82	0.850
12	San Francisco–San Mateo–Redwood City, CA	129	0.730
13	Portland–Vancouver–Beaverton, OR–WA	93	0.678
14	Washington–Arlington–Alexandria, DC–VA–MD–WV	100	0.658
15	Detroit–Livonia–Dearborn, MI	165	0.653
No Comcast			
1	Los Angeles–Long Beach–Glendale, CA	1,233	9.200
2	New York–White Plains–Wayne, NY–NJ	939	4.154
3	Santa Ana–Anaheim–Irvine, CA	144	3.007
4	San Diego–Carlsbad–San Marcos, CA	216	2.926
5	Phoenix–Mesa–Scottsdale, AZ	178	2.041
6	Dallas–Plano–Irving, TX	160	1.976
7	Tampa–St. Petersburg–Clearwater, FL	144	1.556
8	Fort Worth–Arlington, TX	87	1.507
9	Riverside–San Bernardino–Ontario, CA	83	1.371
10	Las Vegas–Paradise, NV	61	1.170
11	San Antonio, TX	94	1.127
12	Columbus, OH	139	0.980
13	Cleveland–Elyria–Mentor, OH	143	0.933
14	Austin–Round Rock, TX	39	0.875
15	Cincinnati–Middletown, OH–KY–IN	127	0.731

Table 1.4:

Descriptive Statistics

This table provides averages and standard deviations of demographic indicators in urban Comcast and no Comcast census tracts. Population, percent living in urban areas, median age, and average household size are obtained from the 2010 Decennial Census. All other demographic variables are calculated using 2007-2011 ACS 5-year estimates. Cost-burdened homeownership captures the fraction of homeowners paying 30 percent or more of income on housing-related payments, as defined by the U.S. Department of Housing and Urban Development (HUD). Bank branch access is measured as the number of full-service bank branches located within 2 miles of a census tract's population centroid as of 2010 using data from the FDIC. Data on broadband connections are obtained from the FCC's Form 477 as of December 2011. Broadband is defined as fixed internet connections with minimum download speeds of 3 Mbps. Means and standard deviations are weighted by each census tract's 2010 population. Statistics for variables other than population, median age, average household size, and number of bank branches are reported in percent. Column 5 reports t-statistics from the Welch two sample test of difference in means.

	Comcast ($N = 2,430$)		**No Comcast** ($N = 2,826$)		
	Mean	SD	Mean	SD	Diff.
	(1)	(2)	(3)	(4)	(5)
Population (2010)	8845.55	7372.68	11487.21	9674.52	11.21
Annual income under $35,000	29.82	12.73	30.03	12.01	0.61
Annual income $35,000 – $50,000	13.42	4.28	13.83	4.04	3.52
Living in urban areas (2010)	98.80	6.49	97.44	10.33	-5.78
Median age (2010)	36.61	5.15	36.68	5.89	0.42
Average household size (2010)	2.70	0.49	2.82	0.61	7.66
Owner-occupancy					
Annual income under $35,000	44.70	20.16	44.08	19.08	-1.13
Annual income $35,000 – $50,000	56.86	20.90	55.18	19.93	-2.98
With school-aged child	30.65	11.06	30.42	11.24	-0.76
Cost-burdened homeowners	42.41	11.56	43.56	11.81	3.55
Employment rate	90.44	4.62	90.98	3.71	4.59
High school diploma or higher	89.91	9.33	89.31	9.80	-2.26
Number of bank branches	18.71	24.37	11.72	11.19	-13.01
Broadband connections	47.33	15.81	36.90	18.90	-21.77

Table 1.5:

Mortgage Characteristics by Comcast Coverage

This table provides average levels of key variables relating to home ownership for urban Comcast and no Comcast census tracts. *'04-'08 purchase* refers to statistics for home purchase mortgages originated between 2004 and 2008, while *'08-'11 refinance* reports the same averages for refinance mortgages originated between 2008 and 2011. All households refer to the universe of purchase and refinance originations for the respective periods. Eligible households have income below 185 percent of the FPL for a three-person family, and ineligible households have income between 185 percent of the FPL for five- and six-person families. All variables, with the exception of interest rates, debt-to-income, combined loan-to-value, and credit scores, are calculated using the universe of HMDA entries for conventional, one- to four-family, owner-occupied fixed rate mortgages. The remaining variables are computed using a matched data set of HMDA and GSE loan performance files, and comprise a subset of originated loans that were sold to Fannie Mae and Freddie Mac. ***, **, and * represent statistical significance of the Welch two sample t-test between means of each group across Comcast and No Comcast census tracts, at the 1%, 5%, and 10% level.

	Comcast ($N = 2,430$)			**No Comcast** ($N = 2,826$)		
	All	Eligible	Ineligible	All	Eligible	Ineligible
	(1)	(2)	(3)	(4)	(5)	(6)
HH income ($ thousands)						
'04-'08 purchase	98.75	24.00	45.71	113.41***	23.90	45.65
'08-'11 refinance	99.94	24.77	50.21	104.21***	24.81	50.21
Loan count						
'04-'08 purchase	646.16	30.15	49.34	661.30	38.49***	51.84***
'08-'11 refinance	389.01	20.09	21.51	324.44***	21.55***	21.17
Loan amount ($ thousands)						
'04-'08 purchase	233.50	114.08	135.74	291.05***	117.84**	139.00***
'08-'11 refinance	203.13	122.83	148.46	245.20***	136.01***	166.76***
Interest rate (percent)						
'04-'08 purchase	6.06	6.18	6.09	6.06	6.18	6.09
'08-'11 refinance	4.98	4.88	4.89	4.98	4.86	4.88
Debt-to-income						
'04-'08 purchase	36.49	36.46	37.54	37.37***	36.44	37.50
'08-'11 refinance	31.81	31.92	32.47	33.46***	33.03***	33.00**
Combined LTV						
'04-'08 purchase	80.32	77.37	80.68	78.19***	76.65	77.71***
'08-'11 refinance	67.05	58.97	63.60	63.76***	57.67**	61.93***
Credit score						
'04-'08 purchase	734.61	726.86	732.75	735.59**	726.63	733.93
'08-'11 refinance	753.31	756.39	756.67	752.94	756.80	757.57
Male (percent)						
'04-'08 purchase	59.29	45.46	53.26	60.09***	46.52**	54.39***
'08-'11 refinance	57.17	40.42	50.54	58.50***	43.21***	53.52***
Black (percent)						
'04-'08 purchase	19.05	20.10	20.89	14.87***	15.41***	15.53***
'08-'11 refinance	18.63	18.93	18.67	14.82***	14.56***	14.40***
Hispanic (percent)						
'04-'08 purchase	12.88	14.45	13.23	20.27***	20.39***	18.60***
'08-'11 refinance	9.61	11.53	11.00	17.12***	19.94***	18.87***

Table 1.6:

Broadband Access and Refinancing Activity

This table reports the effect of Internet Essentials on refinancing outcomes. I estimate the following triple differences regression at the eligibility group level (columns 1 to 3) and loan level (column 4):

$$\begin{aligned} y_{i,c,t} =& \alpha + \beta(Eligible_{i,c,t} \times Comcast_c \times Post_t) + X'_{i,c,t}\Phi + \rho_1(\lambda_t \times \gamma_c) + \rho_2(Eligible_{i,c,t} \times \lambda_t) \\ &+ \rho_3(Eligible_{i,c,t} \times \gamma_c) + \epsilon_{i,c,t}. \end{aligned}$$

Dependent variables in columns 1 and 2 are the annual number of refinance mortgage originations and applications for each eligibility group, respectively. In column 3, denial rates are measured as the ratio of refinance applications denied by financial institutions to total applications. The dependent variable in column 4 is the interest rate for an originated refinance loan. The sample consists of all loan applications from 2008 to 2015 in urban central metro counties. $Eligible_{i,c,t}$ is an indicator for whether a refinance mortgage is associated with a household that qualifies for Internet Essentials based on annual income. $Comcast_c$ is a continuous measure for the fraction of a census tract's population with Comcast access and $Post_t$ is an indicator for post-Internet Essentials launch in 2012. Columns 1 and 2 report PPML results and columns 3 and 4 report OLS results. Group means are reported as of 2011, the last pre-treatment year. I include average income and loan amount as eligibility group controls, and income, loan amount, race, sex, number of applicants, combined LTV, DTI, credit score, and maturity as loan characteristics controls. All specifications incorporate eligibility-year, eligibility-census tract, and census tract-year fixed effects. Robust standard errors reported in parentheses are clustered by Public Use Microdata Area (PUMA). ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable	Number of originations (1)	Number of applications (2)	Denial rate (3)	Interest rate (4)
$(Eligible_{i,c,t} \times Comcast_c \times Post_t)$	0.060***	0.060**	0.002	0.003
	(0.025)	(0.020)	(0.007)	(0.014)
Mean of dependent variable				
Eligible	6.48	14.02	40.86	4.35
Ineligible	6.09	10.96	30.95	4.28
Controls				
Eligibility group	✓	✓	✓	
Loan characteristics				✓
Fixed effects	✓	✓	✓	✓
Observations	81,782	82,768	82,768	115,662
Adjusted R^2	0.64	0.72	0.22	0.86

Table 1.7:

Broadband Access and Mortgage Costs

This table reports the effect of Internet Essentials on mortgage costs. I estimate the following triple differences regression at the household level:

$$
\begin{aligned}
m_{i,p,t} =& \alpha + \beta(Eligible_{i,p,t} \times Comcast_p \times Post_t) + Z'_{i,p,t}\Phi + \rho_1(\lambda_t \times \gamma_p) + \rho_2(Eligible_{i,p,t} \times \lambda_t) \\
&+ \rho_3(Eligible_{i,p,t} \times \gamma_p) + \epsilon_{i,p,t}.
\end{aligned}
$$

Dependent variables $m_{i,p,t}$ are the natural logarithm of monthly mortgage payments (column 1) and the mortgage to income ratio (column 2). The sample consists of all ACS respondents from 2008 to 2015 in metropolitan PUMAs. I restrict the sample to households that have a mortgage and lived in the current home for at least three years. $Eligible_{i,p,t}$ is an indicator for Internet Essentials eligibility. $Comcast_p$ is an indicator for Comcast access (over 90 percent coverage is treated, less than 10 percent coverage is control). $Post_t$ is an indicator for post-Internet Essentials launch in 2012. Panel A employs the full control group of low-income and higher income ineligibles. Panel B only uses low-income ineligibles as the control group. All specifications report OLS results and group means ($ thousands and percent) are reported as of 2011, the last pre-treatment year. Household controls include age, age-squared, sex, marriage status, number of children, employment status, value of house (log), income (log), years since household moved to area, indicator for taxes included in mortgage payments, poverty status, and the Hauser and Warren Socioeconomic Index. All specifications incorporate eligibility-year, eligibility-PUMA, and PUMA-year fixed effects. Robust standard errors reported in parentheses are clustered by PUMA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable	Log(mortgage payment)	Mortgage to income
	(1)	(2)
Panel A: All ineligibles control group		
($Eligible_{i,p,t} \times Comcast_p \times Post_t$)	−0.025**	−0.015***
	(0.011)	(0.004)
Mean of dependent variable		
Eligible	0.82	36.98
Ineligible	0.83	28.29
Household controls	✓	✓
Fixed effects	✓	✓
Observations	385,122	385,122
Adjusted R^2	0.51	0.57
Panel B: Low-income ineligibles control group		
($Eligible_{i,p,t} \times Comcast_p \times Post_t$)	−0.038**	−0.014***
	(0.015)	(0.005)
Mean of dependent variable		
Eligible	0.82	36.98
Ineligible	0.66	37.99
Household controls	✓	✓
Fixed effects	✓	✓
Observations	182,900	182,900
Adjusted R^2	0.49	0.52

Table 1.8:

Heterogeneous Effects by Bank Branch Access

This table reports the heterogeneous effects of Internet Essentials based on bank branch access. I estimate the following triple differences regression at the eligibility group level:

$$y_{i,c,t} = \alpha + \beta(Eligible_{i,c,t} \times Comcast_c \times Post_t) + X'_{i,c,t}\Phi + \rho_1(\lambda_t \times \gamma_c) + \rho_2(Eligible_{i,c,t} \times \lambda_t) + \rho_3(Eligible_{i,c,t} \times \gamma_c) + \epsilon_{i,c,t}.$$

Dependent variables $y_{i,c,t}$ are the fraction of refinances mortgages originated by fintech lenders (column 1) and the number of originations (column 2). The sample consists of all originated mortgages from 2008 to 2015 in urban central metro counties. $Eligible_{i,c,t}$ is an indicator for whether a refinance mortgage is associated with a household that qualifies for Internet Essentials based on annual income. $Comcast_c$ is a continuous measure for the fraction of a census tract's population with Comcast access and $Post_t$ is an indicator for post-Internet Essentials launch in 2012. Bank branch access is defined as the number of full-service branch locations within a 2 mile radius of a census tract's population center. I classify census tracts as low (bottom quintile), mid (third quintile), and high (top quintile) based on bank branch access. Group means (percent and loan count) are reported as of 2011, the last pre-treatment year. Column 1 reports OLS results and columns 2 through 4 report PPML results. All specifications include controls for average income and loan amount as well as eligibility-year, eligibility-census tract, and census tract-year fixed effects. Robust standard errors reported in parentheses are clustered by Public Use Microdata Area (PUMA). ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

		Originations		
Dependent variable	% Fintech	Low	Mid	High
	(1)	(2)	(3)	(4)
$(Eligible_{i,c,t} \times Comcast_c \times Post_t)$	-0.006	0.091***	0.050	-0.021
	(0.009)	(0.035)	(0.036)	(0.068)
Number of bank branches < 2 mi		4.12	14.97	57.92
Mean of dependent variable				
Eligible	4.30	7.08	6.34	3.93
Ineligible	7.16	6.70	5.82	3.82
Eligibility group controls	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓
Observations	72,578	22,626	18,608	4,872
Adjusted R^2	0.11	0.67	0.62	0.43

Table 1.9:

Heterogeneous Effects by Educational Attainment

This table reports the heterogeneous effects of Internet Essentials relating to income and educational attainment. Dependent variables are log annual household income (column 1), number of refinance originations by group (columns 2, 3, and 4), and log monthly mortgage payments (columns 5, 6, 7, and 8). The sample is from 2008 to 2015. $Eligible_{i,c,t}$ ($Eligible_{i,p,t}$) is an indicator for whether an eligibility group (household) qualifies for Internet Essentials. $Comcast_c$ is a continuous measure of Comcast coverage in census tract c and $Comcast_p$ is a binary indicator for Comcast availability in PUMA p. $Post_t$ is an indicator for post-Internet Essentials launch in 2012. In columns 2, 3, and 4, I subset census tracts by educational attainment using the fraction of the population with at least a high school diploma (ACS 2007-2011 5-year estimates). Low and high census tracts refer to the bottom and top quartiles, respectively. In columns 5, 6, 7, and 8, I study owner-occupied households (with a mortgage) whose heads' educational attainment is high school diploma or less (low) and at least some college education (high). Columns 5 and 6 incorporate all ineligibles as the control group and columns 7 and 8 only use low-income ineligibles as the control group. Group means of dependent variables ($ thousands and loan counts) are reported as of 2011, the last pre-treatment year. All specifications include controls for average income and loan amount (columns 2, 3, and 4) or household characteristics (columns 1, 5, 6, 7, 8). Eligibility-year, eligibility-census tract/PUMA, and census tract/PUMA-year fixed effects are used depending on the level of geography. Robust standard errors reported in parentheses are clustered by PUMA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable	Log(income)	Originations			Log(mortgage payment) All ineligibles		Log(mortgage payment) Low-income ineligibles	
		Low	Mid	High	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
($Eligible_{i,p,t} \times Comcast_p \times Post_t$)	0.004				-0.054***	0.006	-0.083***	0.012
	(0.012)				(0.015)	(0.013)	(0.020)	(0.020)
($Eligible_{i,c,t} \times Comcast_c \times Post_t$)		0.125**	0.044	0.059**				
		(0.053)	(0.032)	(0.028)				
High school diploma or higher		0.72	0.90	0.97				
Mean of dependent variable								
Eligible	29.91	5.20	6.81	6.99	0.78	0.86	0.78	0.86
Ineligible	24.35	3.51	6.23	7.82	0.74	0.91	0.61	0.73
Controls								
Eligibility group		✓	✓	✓				
Household	✓				✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Observations	102,868	19,958	34,550	27,274	184,172	200,739	98,774	83,484
Adjusted R^2	0.61	0.56	0.65	0.66	0.51	0.50	0.50	0.49

Table 1.10:

Robustness Measures and Sensitivity Analyses

This table provides sensitivity analyses for the effect of Internet Essentials on refinancing. Dependent variables are prepayment indicator (column 1), number of refinance originations by group (columns 2, 3, 4, and 5), and log monthly rent payments (column 6). The sample is restricted to urban metropolitan areas between 2008 and 2015. $Eligible_{i,c,t}$ ($Eligible_{i,p,t}$) is an indicator for whether a loan or eligibility group (household) qualifies for Internet Essentials. $Comcast_c$, $AT\&T_c$, and $Charter_c$ are continuous measures of Comcast, AT&T, and Charter coverage in census tract c, respectively. $Comcast_p$ is a binary indicator for Comcast availability in PUMA p. $Post_t$ is an indicator for post-Internet Essentials launch in 2012. Column 3 replaces the income thresholds for treated (185% FPL for five- and six-person family) and control (185% FPL for seven- and eight-person family) groups. Group means of dependent variables (percent, loan counts, and $ thousands) are reported as of 2011, the last pre-treatment year. All specifications include controls for loan characteristics at origination (column 1), average income and loan amount (columns 2, 3, 4, and 5), and household characteristics (column 6). Eligibility-year, eligibility-census tract/PUMA, and census tract/PUMA-year fixed effects are used depending on the level of geography. Robust standard errors reported in parentheses are clustered by PUMA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

		Originations				
Dependent variable	Prepayment (1)	Baseline (2)	Placebo cutoff (3)	Placebo ISP (4)	(5)	Log(rent payment) (6)
($Eligible_{i,c,t} \times Comcast_c \times Post_t$)	0.033**	0.060**	0.003			
	(0.014)	(0.025)	(0.011)			
($Eligible_{i,c,t} \times AT\&T_c \times Post_t$)				-0.009		
				(0.023)		
($Eligible_{i,c,t} \times Charter_c \times Post_t$)					-0.009	
					(0.045)	
($Eligible_{i,p,t} \times Comcast_p \times Post_t$)						0.006
						(0.013)
Mean of dependent variable						
Eligible	41.99	6.48	6.46	6.48	6.48	0.74
Ineligible	59.76	6.09	6.79	6.09	6.09	0.77
Controls						
Loan characteristics	✓					
Eligibility group		✓	✓	✓	✓	
Household						✓
Fixed effects	✓	✓	✓	✓	✓	✓
Observations	220,898	81,782	76,972	81,782	81,782	238,188
Adjusted R^2	0.21	0.64	0.64	0.64	0.64	0.42

Table 1.11:

Falsification Tests for Likelihood of Program Access

This table studies the effect of Internet Essentials in census tracts that are more or less likely to be impacted by the program. The dependent variable in all specifications is the number of refinance originations by eligibility group. The sample is restricted to urban metropolitan areas between 2008 and 2015. $Eligible_{i,c,t}$ is an indicator for whether group i qualifies for Internet Essentials. $Comcast_c$ is a continuous measures of Comcast coverage in census tract c. $Post_t$ is an indicator for post-Internet Essentials launch in 2012. Column 1 is the baseline specification that includes all census tracts. Columns 2 to 4 subset the census tracts into the bottom quartile (low), top quartile (high), and the 25th to 75th percentile (mid) by the fraction of households in owner-occupied dwellings with at least one child under the age of 18 as of 2011. Columns 5 to 7 subset census tracts by the fraction of cost-burdened homeowners (paying 30 percent or more of income on housing costs) as of 2011. Columns 8 to 10 subset census tracts by the fraction of low-income homeowners with a school-aged child that reported having a high-speed broadband connection. This variable is obtained from the 2013 ACS 1-year microdata and assignment is at the PUMA level. Means of sorting variables are reported in percent. Group means are reported as of 2011, the last pre-treatment year. All specifications include controls for household characteristics. Eligibility-year, eligibility-census tract, and census tract-year fixed effects are used depending on the level of geography. Robust standard errors reported in parentheses are clustered by PUMA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

		School-aged children			Cost-burdened homeowners			Broadband subscription rate		
Dependent variable: Originations	Baseline (1)	Low (2)	Mid (3)	High (4)	Low (5)	Mid (6)	High (7)	Low (8)	Mid (9)	High (10)
$(Eligible_{i,c,t} \times Comcast_c \times Post_t)$	0.060***	0.007	0.078***	0.076***	0.024	0.054	0.163***	0.049	0.038	0.092***
	(0.020)	(0.040)	(0.030)	(0.021)	(0.033)	(0.039)	(0.047)	(0.044)	(0.037)	(0.035)
Mean of sorting variable		19.64	30.82	44.55	30.10	45.10	61.07	74.36	85.40	94.05
Mean of dependent variable										
Eligible	6.48	5.42	6.65	7.08	6.71	6.79	5.57	5.56	7.02	5.98
Ineligible	6.09	4.68	6.38	6.84	6.84	6.26	4.62	5.36	6.47	5.89
Eligibility group controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	81,782	21,244	33,340	27,198	31,440	31,614	19,018	23,414	36,444	21,924
Adjusted R^2	0.64	0.59	0.64	0.67	0.65	0.65	0.59	0.61	0.66	0.63

1.B Bibliography

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao (2017) "Systematic mistakes in the mortgage market and lack of financial sophistication," *Journal of Financial Economics*, 123 (1), 42–58.

Agarwal, Sumit, Souphala Chomsisengphet, Hua Kiefer, Leonard C. Kiefer, and Paolina C. Medina (2020) "Inequality During the COVID-19 Pandemic: The Case of Savings from Mortgage Refinancing," *SSRN Electronic Journal*, 10.2139/ssrn.3750133.

Agarwal, Sumit, John C. Driscoll, and David I. Laibson (2013) "Optimal Mortgage Refinancing: A Closed-Form Solution," *Journal of Money, Credit, and Banking*, 45 (4), 591–622.

Agarwal, Sumit, Richard J. Rosen, and Vincent Yao (2016) "Why Do Borrowers Make Mortgage Refinancing Mistakes?" *Management Science*, 62 (12), 3494–3509.

Andersen, Steffen, John Y. Campbell, Kasper M. Nielsen, and Tarun Ramadorai (2020) "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market," *American Economic Review*, 110 (10), 3184–3230.

Argyle, Bronson, Taylor D Nadauld, and Christopher Palmer (2020) "Real Effects of Search Frictions in Consumer Credit Markets," Working Paper 26645, National Bureau of Economic Research.

Bajo, Emanuele and Massimiliano Barbi (2018) "Financial illiteracy and mortgage refinancing decisions," *Journal of Banking & Finance*, 94, 279–296.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace (2022) "Consumer-lending discrimination in the FinTech Era," *Journal of Financial Economics*, 143 (1), 30–56.

Bhutta, Neil and Lisa Dettling (2018) "Money in the Bank? Assessing Families' Liquid Savings using the Survey of Consumer Finances," *FEDS Notes*, 2018 (2275).

Brown, James R., J. Anthony Cookson, and Rawley Z. Heimer (2019) "Growing up without finance," *Journal of Financial Economics*, 134 (3), 591–616.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru (2018) "Fintech, regulatory arbitrage, and the rise of shadow banks," *Journal of Financial Economics*, 130 (3), 453–483.

Byrne, Shane, Kenneth Decine, Michael King, Yvonne McCarthy, and Christopher Palmer (2022) "The Last Mile of Monetary Policy: Consumer Inattention, Disclosures, and the Refinancing Channel," Working Paper.

Campbell, John Y. (2006) "Household Finance," *The Journal of Finance*, 61 (4), 1553–1604.

Célerier, Claire and Adrien Matray (2019) "Bank-Branch Supply, Financial Inclusion, and Wealth Accumulation," *The Review of Financial Studies*, 32 (12), 4767–4809.

Coleman-Jensen, Alisha, Matthew P. Rabbitt, Christian A. Gregory, and Anita Singh (2016) "Household Food Security in the United States in 2015," *USDA Economic Research Report* (215).

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin (2019) "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects," https://arxiv.org/abs/1903.01690.

Covas, Francisco (2019) "Some Facts About Bank Branches and LMI Customers," *Bank Policy Institute Reports and Research Notes*, https://bpi.com/some-facts-about-bank-branches-and-lmi-customers/.

Defusco, Anthony A. and John Mondragon (2020) "No Job, No Money, No Refi: Frictions to Refinancing in a Recession," *The Journal of Finance*, 75 (5), 2327–2376.

Deng, Yongheng and John M. Quigley (2012) "Woodhead Behavior and the Pricing of Residential Mortgages," *Working Paper IRES2012-025, NUS Institute of Real Estate Studies*.

Di Maggio, Marco, Dimuthu Ratnadiwakara, and Don Carmichael (2021) "Invisible Primes: Fintech Lending with Alternative Data," *SSRN Electronic Journal*.

Fonseca, Julia and Adrien Matray (2022) "The Real Effects of Banking the Poor: Evidence from Brazil," Working Paper 30057, National Bureau of Economic Research.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery (2019) "The Role of Technology in Mortgage Lending," *The Review of Financial Studies*, 32 (5), 1854–1899.

Gerardi, Kristopher, Lauren Lambie-Hanson, and Paul Willen (2021) "Racial Differences in Mortgage Refinancing, Distress, and Housing Wealth Accumulation During COVID-19," jun, 10.21799/frbp.dp.2021.02.

Gerardi, Kristopher, Paul Willen, and David Hao Zhang (2020) "Mortgage Prepayment, Race, and Monetary Policy," oct, 10.29412/res.wp.2020.07.

Goodstein, Ryan M. (2013) "Refinancing Trends among Lower Income and Minority Homeowners during the Housing Boom and Bust," *Real Estate Economics*, 42 (3), 690–723.

Gourieroux, C., A. Monfort, and A. Trognon (1984) "Pseudo Maximum Likelihood Methods: Applications to Poisson Models," *Econometrica*, 52 (3), 701–720.

Gruber, Jonathan (1994) “The incidence of mandated maternity benefits,” *American Economic Review*, 622–641.

Haughwout, Andrew F., Donghoon Lee, Joelle Scally, and Wilbert van der Klaauw (2021) “Mortgage Rates Decline and (Prime) Households Take Advantage,” *Federal Reserve Bank of New York Liberty Street Economics*.

Horrigan, John B. (2014) “The Essentials of Connectivity: Comcast’s Internet Essentials Program and a Playbook for Expanding Broadband Adoption and Use in America,” *Technical Report*.

——— (2019) “Reaching the Unconnected: Benefits For Kids and Schoolwork Drive Broadband Subscriptions, but Digital Skills Training Opens Doors to Household Internet Use for Jobs and Learning,” *Technical Report*.

Horrigan, John B. and Maeve Duggan (2015) “Home Broadband 2015,” *Pew Research Center*, `http://www.pewinternet.org/2015/12/21/2015/Home-Broadband-2015/`.

Hussain, Hibah, Nick Russo, Danielle Kehl, and Patrick Lucey (2013) “The Cost of Connectivity 2013,” *New America Policy Paper*, `https://www.newamerica.org/oti/policy-papers/the-cost-of-connectivity-2013/`.

Jiang, Erica Xuewei, Gloria Yang Yu, and Jinyuan Zhang (2022) “Bank Competition amid Digital Disruption: Implications for Financial Inclusion,” *SSRN Electronic Journal*.

Johnson, Eric J, Stephan Meier, and Olivier Toubia (2018) “What’s the Catch? Suspicion of Bank Motives and Sluggish Refinancing,” *The Review of Financial Studies*, 32 (2), 467–495.

Jung, Sakong and Alexander K. Zentefis (2022) “Bank Access Across America,” Working Paper.

Keys, Benjamin J., Devin G. Pope, and Jaren C. Pope (2016) “Failure to Refinance,” *Journal of Financial Economics*, 122 (3), 482–499.

Knutson, Ryan (2015) “How Fast Internet Affects Home Prices,” *Wall Street Journal*, `https://www.wsj.com/articles/how-fast-internet-affects-home-pricesfast-internet-access-can-boost-a-homes-value-1435694077`.

Longstaff, Francis A. (2005) “Borrower Credit and the Valuation of Mortgage-Backed Securities,” *Real Estate Economics*, 33 (4), 619–661.

Maturana, Gonzalo and Jordan Nickerson (2018) “Teachers Teaching Teachers: The Role of Workplace Peer Effects in Financial Decisions,” *The Review of Financial Studies*, 32 (10), 3920–3957.

Philippon, Thomas (2016) “The FinTech Opportunity,” Working Paper 22476, National Bureau of Economic Research.

Schwartz, Eduardo S. and Walter N. Torous (1989) “Prepayment and the Valuation of Mortgage-Backed Securities,” *The Journal of Finance*, 44 (2), 375–392.

Silva, J. M. C. Santos and Silvana Tenreyro (2006) “The Log of Gravity,” *The Review of Economics and Statistics*, 88 (4), 641–658.

Stanton, Richard (1995) “Rational Prepayment and the Valuation of Mortgage-Backed Securities,” *Review of Financial Studies*, 8 (3), 677–708.

Wolff, Edward N. (2016) “Household Wealth Trends in the United States, 1962 to 2013: What Happened over the Great Recession?” *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2 (6), 24–43.

Yang, Keer (2021) “Trust As an Entry Barrier: Evidence from FinTech Adoption,” *SSRN Electronic Journal*.

Yogo, Motohiro, Andrew Whitten, and Natalie Cox (2021) "Financial Inclusion Across the United States," *SSRN Electronic Journal*.

Zuo, George W. (2021) "Wired and Hired: Employment Effects of Subsidized Broadband Internet for Low-Income Americans," *American Economic Journal: Economic Policy*, 13 (3), 447–482.

CHAPTER 2

Intangible Value

(with Andrea L. Eisfeldt and Dimitris Papanikolaou)

Value investing requires a fundamental anchor in order to determine which stocks are priced "expensively" vs. "cheaply" relative to their fundamental value. Using the book value of a firm's assets as the value anchor was popularized by Fama and French (1992, 1993), and the value effect subsequently became one of the most storied and studied anomalies in finance. However, the value factor has underperformed for at least a decade.[1] We argue that one driver of value's poor performance during this period is the deteriorating quality of book assets as a fundamental anchor due to the omission of intangible assets. Correctly defining the fundamental anchor for the value factor is important both in the context of rational explanations of value, in which book assets capture assets in place, and for behavioral explanations, in which market to book ratios represent a measure of mispricing.

Intangible assets have become an important and fast-growing part of firms' capital stocks. Corrado et al. (2009) estimated intangibles to be about one third of the US non-residential capital stock in 2003, while, using more recent data, Eisfeldt and Papanikolaou (2013b), Falato et al. (2013), Belo et al. (2019), and Ewens et al. (2020) all estimate the contribution of intangible capital to overall corporate capital stocks to be around one half. In addition, these same studies report much higher investment rates for intangible assets relative to physical assets. The majority of intangible

[1] See, for example, Figure 7.6 in Ang (2014). We independently document the decline in value below.

assets are created by investments in employee, brand, and knowledge capital that are expensed and thus do not appear on corporate balance sheets. This has resulted in a growing mis-measurement of book assets.

We propose an intangible-augmented value factor ("intangible value", HML^{INT}) and construct it using a very simple modification to the standard Fama and French value factor (HML^{FF}). Our construction of HML^{INT} precisely follows the Fama and French methodology. The key difference is that we add intangible assets to the book equity of each firm, which is widely used as the traditional value anchor.[2] We also perform our intangible value sort within industries, which is useful for two reasons. First, as documented by Asness et al. (2000) and confirmed in our data, both traditional and intangible value are primarily within-industry phenomena. Measuring value within industries thus increases efficiency and reduces exposure to unpriced risk. Daniel et al. (2020) document the large increase in Sharpe ratios that can be achieved by reducing exposures to unpriced risks. Second, because accounting practices vary across industries, sorting within industries alleviates some of the criticisms levied at incorporating intangibles into value measures raised by Rizova and Saito (2020). In the Online Appendix, we show that a small (but not negligible) part of the improvement to traditional value arises from sorting firms within industries when constructing intangible value. For ease of comparison with the existing literature on traditional value, we use the standard value factor from the Fama and French data library as the traditional value factor in our study.[3]

We follow the method introduced in Eisfeldt and Papanikolaou (2013b) to measure firm-level stocks of intangible assets. Specifically, we apply the perpetual inventory method to flows of Selling, General, and Administrative (SG&A) expenses, given as-

[2] Note that this implies an inherent assumption that all intangibles are equity backed, which is consistent with, for example, Rampini and Viswanathan (2013) and Falato et al. (2013).

[3] See `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html`.

sumptions about depreciation and initial values. Eisfeldt and Papanikolaou (2013b) build on two seminal contributions in measuring intangible assets. Corrado et al. (2009) use aggregated expenditure data and the perpetual inventory method to estimate the value of three main categories of intangibles: computerized information, R&D, and economic competencies.[4] Lev and Radhakrishnan (2005) document that firms with larger SG&A expenses exhibit greater Solow (1957) residuals. Eisfeldt and Papanikolaou (2013b) extend this work and are the first to construct and analyze firm-level stocks of intangible assets measured as accumulated SG&A expenses. That paper shows that firms with higher stocks of intangible assets outperform firms with lower intangibles, and provides additional evidence supporting the use of SG&A as a measure of intangible investment.[5] Measures of intangible assets using accumulated SG&A are also supported by the subsequent findings in Eisfeldt and Papanikolaou (2014), Zhang (2014), Falato et al. (2013), and Peters and Taylor (2017).

Importantly, we follow Eisfeldt and Papanikolaou (2013b) and include all of SG&A as investment in intangibles, rather than using the subsequent method introduced by Peters and Taylor (2017).[6] That method generally follows Eisfeldt and Papanikolaou (2013b) but uses only 30% of (SG&A minus R&D) plus 100% of R&D as investment in intangibles. There are two important reasons why we do not use the Peters and Taylor (2017) method to construct the stock of intangible assets. The first reason is that the 30% fraction used in Peters and Taylor (2017) is a calibrated number based on a small number of firms many decades ago. Indeed, later work by one of the authors of Peters and Taylor (2017) questions this assumption and attempts to

[4] See also the precursor to that paper, Corrado et al. (2005), for further details.

[5] In particular, firms with more intangible assets using their measure are more productive, smaller, have higher Tobin's Q, executive compensation, and managerial quality scores according to the measure of Bloom and Van Reenen (2007), spend more on information technology (IT), and are more likely to list "loss of key personnel" as a risk factor in their 10-K filings. See also Lev (2000) and Eisfeldt and Papanikolaou (2013a) for further evidence supporting SG&A as intangible investment.

[6] See also Amenc et al. (2020) and Arnott et al. (2021) for studies that use the Peters and Taylor (2017) method to construct intangible capital in the context of value strategies.

construct industry-specific investment ratios for SG&A. Ewens et al. (2020) state that “the only estimate of γ_S (the fraction of SG&A that is intangible investment) comes from Hulten and Hao (2008), who estimate it based on descriptions of income statement items from six pharmaceutical firms in 2006, applying the investment share of expensed items from Corrado, Hulten, and Sichel (2006).”[7]

A second key rationale for using 100% of SG&A to construct intangible capital stocks is that there is no compelling reason to break out R&D expenses but not advertising expenses or other intangible asset expenditures. We argue that including all of SG&A and sorting within industries provides more reliable intangible capital estimates. Similar to advertising expenses, R&D is reported separately by only a subset of firms. As documented by Koh and Reeb (2015), missing values for R&D should not be interpreted as zeros.[8] Without better estimates of the fraction of SG&A spending that is investment in intangible assets, we argue that it is best to use 100% of SG&A and to sort on relative intangible capital stocks across firms that are likely to share accounting practices (i.e., within industries) to avoid introducing noise. Our method reduces reliance on imprecise estimates of free parameters. In addition, using 100% of SG&A better accounts for organization, brand and customer capital, the importance of which can be substantial in many industries. Note that because we sort firms within industries to construct our value factor, any heterogeneity in the fraction of SG&A spending that is investment in intangibles across industries cancels out.[9] This is important because, as we document, accounting practices for allocating costs to SG&A vs. Cost of Goods Sold (COGS) vary systematically across industries.

[7] Note that the latter paper is published as Corrado et al. (2009) and covers a broad set of industries. However, the 30% estimate in Hulten and Hao (2008) is derived from pharmaceutical firms.

[8] See also the related older work by Bound et al. (1982) whose Table 2.2 shows larger differences across industries in R&D spending reported in the National Science Foundation R&D survey (see `https://www.nsf.gov/statistics/industry/`) than in Compustat data.

[9] See the new study by Lev and Srivastava (2019) which makes progress on understanding firm-level variation in the effect of SG&A spending on intangibles.

Our intangible value factor, HML^{INT}, has the following features: (1) It is highly correlated with the traditional value factor, HML^{FF} (76%). (2) It prices standard test assets with lower pricing errors than HML^{FF}. (3) It substantially and significantly outperforms HML^{FF}. The average returns to a portfolio that is long HML^{INT} and short HML^{FF} are 2.11% annually, with a standard deviation of only 6.53%. This long-short portfolio's Sharpe ratio (or equivalently, HML^{INT}'s information ratio with respect to HML^{FF}) is 0.32 over the full sample and 0.62 in data since 2007. This outperformance holds over the entire sample, and is in fact more pronounced in the post-crisis era in which the returns to traditional value have been particularly disappointing. Thus, although HML^{INT} is highly correlated with the original value factor, it has enough independent variation to permit substantial outperformance. The R^2 in a regression of HML^{INT} on HML^{FF} is 58%. The alpha of intangible value in a single traditional value factor model is 3.86% and highly statistically significant.

We examine in detail the potential drivers of intangible value's ability to price standard test assets as well as the traditional value factor and its substantial outperformance. We also decompose the intangible value factor into traditional value and two factors that better isolate the effects of intangible capital. The first is an isolated intangible value factor, HML^{IME}, which sorts firms based only on our measure of the book value of intangible capital relative to the market value of equity. The second decomposition, HML^{UINT}, is constructed by taking long positions in firms that are uniquely in the long leg of HML^{INT} (specifically, not in the long leg of HML^{FF}), and short positions in firms that are uniquely in the short leg. These more isolated measures of intangible value continue to price standard test assets as well as or better than traditional value. The HML^{IME} portfolio has positive and significant alphas in the three- and five-factor models plus momentum, and the HML^{UINT} portfolio has a positive and significant alpha in the five-factor model plus momentum.

We also document important differences in characteristics between firms in the long (and short) legs of intangible and traditional value. It appears that intangible value is long firms with better fundamentals. The long leg of intangible value contains firms with higher productivity, higher earnings to price ratios (thus better valuation metrics by non-book measures), higher profits to assets, and lower debt to earnings. By contrast, traditional value is long firms with lower gross profitability to total assets, lower sales to stockholders' equity, lower sales to book assets, and higher debt to earnings.

Our findings have several implications. First, asset pricing researchers should consider correcting book equity for intangibles as intangible assets are a large and growing part of the corporate capital stock and there is a small gain in model fit from replacing the traditional value factor with the intangible-augmented factor. Second, asset managers should consider using the intangible value factor when implementing a value tilt in a relative value strategy. HML^{INT} appears to capture the value effect in that it prices standard test portfolios just as well as traditional value, but achieves higher average returns and lower volatility. Finally, an active manager can implement a profitable long-short strategy by going long HML^{INT} and short HML^{FF}.

The paper most closely related to ours is Park (Forthcoming), of which we were made aware upon circulating this paper. Because the two papers developed independently, the methodologies differ somewhat. The theoretical benefits of the two key differences in our methodology, namely sorting within industries and using the Eisfeldt and Papanikolaou (2013b) method for constructing intangible stocks using 100% of SG&A expenses, are detailed further in the next section. Empirically, we show that our method leads to an intangible value factor that has a positive alpha of 2.42% that is significant at the 1% level with respect to the intangible value factor constructed using the Peters and Taylor (2017) method, which also does not sort firms within industries (the method used in Park (Forthcoming)). Our paper also makes substantial new contributions relative to Park (Forthcoming), and in general the two studies are complementary. In particular, we investigate the differences between traditional and intangible value in more detail by studying portfolios sorted on intangible assets only to market equity and portfolios consisting of firms that are uniquely in the long or short leg of intangible value.

Additionally, we examine the how the long and short legs of intangible value contribute to the factor's outperformance, and provide examples of how the intangible value portfolio avoids "value traps" and avoids shorting low book-to-market firms whose book values do not reflect their total capital stock. We also examine the firm-level characteristics of the long and short legs of intangible vs. traditional value, and document the substantial differences in productivity, profitability, price to earnings ratios, and leverage. This paper also documents the difference between the intangible value factor and the organization capital factor in Eisfeldt and Papanikolaou (2013b), which also utilizes the accumulated stock of SG&A expenses to measure intangible (organization) capital. The key difference is that the portfolios in Eisfeldt and Papanikolaou (2013b) are formed using sorts on book organization capital to total book assets, rather than sorts on total book assets to market values of equity. As a result, the intangible value portfolio has low loadings on, and cannot be explained by, returns to the organization capital portfolio.

Our study also more formally examines the outperformance of intangible value relative to traditional value. We construct a strategy that is long intangible value and short traditional value and document the performance statistics for that strategy. We show that intangible value has a statistically significant alpha of 3.82% with respect to a single-factor traditional value model. Despite the high correlation between the two value strategies, this is not a near-arbitrage strategy. The appraisal ratio (alpha relative to the root mean squared pricing error) is 0.91. We also examine subsamples to see when the outperformance arises. In terms of average returns, the outperformance appears to be increasing over time, and is highest in the most recent subsample, post-great financial crisis. This is consistent with the importance of intangible assets continuing to grow. This subsample is also of substantial interest because it is also the prolonged period during which the performance of traditional value has been particularly poor.

Finally, we closely follow the Fama and French methodology for constructing book equity, and for constructing the long and short legs of both the traditional and intangible value portfolios. Before adding intangible capital to book equity, we confirm that we can successfully replicate the Fama and French traditional value factor from their data library. This is crucial, because it is well-known that slight changes in methodology can lead to large differences in replication errors and a vast literature on the value effect in finance utilizes the Fama and French series.

The paper proceeds as follows. In Section 2.1 we describe the data sources and the construction of our intangible value factors. In Section 2.2 we document the high correlation between the traditional value factor and the intangible value factor, and the superior performance of the intangible value factor in pricing standard test portfolios. We conduct several important robustness exercises, including examining intangible value portfolios formed only using intangible assets or only using firms that have a different portfolio assignment than that assigned by the traditional value factor. Then, Section 2.3 documents the outperformance of the intangible value factor, particularly in more recent subsamples. Section 2.4 examines the drivers of the differences between intangible and traditional value, and Section 2.5 concludes.

2.1 The Intangible Value Factor (HML^{INT})

In this section, we provide details on how we construct HML^{INT} and discuss our measurement choices in more detail.

2.1.1 Data and Sample

As our goal is to compare the relative pricing and return performance of the published HML factor and our HML^{INT} factor, we first ensure that our factor construction matches the Fama and French (1992, 1993) data construction methodology as closely as possible. Our replicated series of the published HML factor has a correlation with the original series of 98%.

We use standard accounting data from Compustat and stock price data from the Center for Research in Security Prices (CRSP). We obtain returns data for factors and test assets, as well as 12-Industry classifications, from Ken French's website.[10] The sample period of our main study is 1975 to 2018, and we additionally conduct analyses for sub-periods from 1995 to 2018 (post-internet era) and 2007 to 2018 (post-crisis era).

2.1.2 Constructing the Intangible Value Factor

To construct HML^{INT}, we add intangible assets to book equity. That is, we define total book equity as

$$\text{B}_{it}^{\text{INT}} = \text{B}_{it} - \text{GDWL}_{it} + \text{INT}_{it}, \tag{2.1}$$

where B_{it} is book equity, GDWL_{it} is goodwill, and INT_{it} is intangible assets for firm i at time t. We subtract goodwill in order to reduce the effects of merger activity and to alleviate the associated double counting of intangibles. We use the perpetual inventory method following Eisfeldt and Papanikolaou (2013b), Eisfeldt and Papanikolaou (2013a), and Eisfeldt and Papanikolaou (2014) to calculate INT_{it}.

$$\text{INT}_{it} = (1-\delta)\text{INT}_{it-1} + \text{SG\&A}_{it}. \tag{2.2}$$

[10] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

We initialize $\text{INT}_{i0} = \text{SG\&A}_{i1}/(g + \delta)$ using the observation for SG&A when the firm first appears in Compustat. We set $g = 0.1$, which is approximately the average growth rate for SG&A in our sample, and assume a depreciation rate of $\delta = 0.2$ following Eisfeldt and Papanikolaou (2014). We apply this algorithm to all firms in Compustat from 1950, and begin our main sample in 1975.

Once we have a firm-level measure of B^{INT}, we form $\text{B}^{\text{INT}}/\text{M}$ portfolios in June of each year using book equity values reported in the previous year and market equity values from the previous December. To do this, we sort firms into tercile buckets by $\text{B}^{\text{INT}}/\text{M}$ every period *within each industry.* Following the procedure of Fama and French (1992, 1993), we compute industry HML^{INT} returns using six value-weighted portfolios formed on size and book-to-market. Lastly, we value-weight the industry HML^{INT} returns by each industry's market capitalization. The resulting market-level factor is the primary intangible value factor used throughout the paper.[11]

Our industry-based sorting method is notably distinct from traditional methods popularized by Fama and French and adopted by recent papers in this literature. We argue that an industry-level sort is preferable to constructing an economy-wide sort for several reasons. First, we confirm the findings of Asness et al. (2000) that value has consistently been a within-industry phenomenon, for both traditional and intangible value. As reported in Table 2.1, book-to-market's ability to predict stock returns is almost entirely driven by within-industry variation. Using either the traditional or intangible measure of book-to-market, the across-industry contributions to market-wide value are not significantly different from zero. Additionally, the within-industry T-statistics (8.65 and 8.75, respectively) are actually larger than the market-wide T-statistics (5.82 and 7.56). Measuring value within industries thus reduces noise and exposure to unpriced risk, which should increase achievable Sharpe ratios ((Daniel

[11] Further details on the factor construction methodology we employ can be found in the Online Appendix.

et al., 2020)).

Another important reason for sorting value within industries is to address the readily documented heterogeneity in accounting practices across industries. Koh and Reeb (2015) document the fact that missing R&D should not be interpreted as zeros, arguing that doing so can underestimate intangible capital expenditures for a large subset of firms. Panel A of Table 2.2 documents the variation in the fraction of missing R&D observations across industries, which range from 11% to 99%. The mean and median fraction of missing R&D observations are 54% and 55% respectively. This implies that the majority of R&D data are in fact missing observations. Additionally, whether R&D expenditures are broken out separately from SG&A depends on industry standard practices. Due to the discrepancy in reporting practices for R&D, we argue that sorting within industries and accumulating 100% of SG&A to measure intangible capital is the most reliable method currently available for constructing intangible capital stocks. This method, as opposed to those that accumulate organization (SG&A minus R&D) and knowledge (R&D) capital expenditures separately, avoids setting missing R&D to zero as is commonly done in the literature ((Park, Forthcoming; Peters and Taylor, 2017)). By accumulating 100% of SG&A and sorting firms within industries, we minimize the number of assumed parameters.

Panel B of Table 2.2 documents the variation across industries in the contribution of SG&A to total costs as measured by (SG&A plus COGS). Such variation could lead to industry under- or over-weighting if intangible value sorts are not conducted within industries. Panel C of Table 2.2 confirms the possibility of distorted industry weights by reporting the variation of changes to the book to market ratio when intangibles are included. While the purpose of including intangibles is in fact to modify B/M, we argue the most reliable estimates thus far are those done on a relative basis between firms in the same industry using 100% of SG&A.

2.1.3 Additional Intangible Value Factors

We construct various alternative measures of intangible value in order to analyze the unique pricing ability of $\mathrm{HML}^{\mathrm{INT}}$ and ensure the robustness of our main results.

In terms of alternative long-short hedged portfolios, $\mathrm{HML}^{\mathrm{IME}}$ is a value factor that sorts firms into high and low buckets based on intangible assets-to-market equity, or INT/M, instead of $\mathrm{B}^{\mathrm{INT}}/\mathrm{M}$. Moreover, $\mathrm{HML}^{\mathrm{UINT}}$ sorts firms on $\mathrm{B}^{\mathrm{INT}}/\mathrm{M}$ but only takes long positions on firms that are *uniquely* in the long leg of $\mathrm{HML}^{\mathrm{INT}}$ (i.e., not sorted in the long leg of $\mathrm{HML}^{\mathrm{FF}}$), and short positions on firms that are *uniquely* in the short leg of $\mathrm{HML}^{\mathrm{INT}}$ (i.e., not sorted in the short leg of $\mathrm{HML}^{\mathrm{FF}}$). Lastly, INT-FF is a factor that is long $\mathrm{HML}^{\mathrm{INT}}$ and short $\mathrm{HML}^{\mathrm{FF}}$, and IME-FF is long $\mathrm{HML}^{\mathrm{IME}}$ and short $\mathrm{HML}^{\mathrm{FF}}$. For these two factors, there may be firms sorted into the same long-short legs but with different portfolio weights.

In the Online Appendix, we construct alternative versions of $\mathrm{HML}^{\mathrm{INT}}$ and $\mathrm{HML}^{\mathrm{FF}}$ to examine the robustness of our results on pricing and outperformance. First, we compare our $\mathrm{HML}^{\mathrm{INT}}$ to $\mathrm{HML}^{\mathrm{INDFF}}$, which is the traditional value factor that follows our within-industry sorting and weighting methodology. Similarly, we analyze the performance of $\mathrm{HML}^{\mathrm{INT}}$ that drops financials (SIC codes 6,000-6,999), regulated utilities (4,900-4,999), and firms categorized as public service, international affairs, or non-operating establishments (9,000+), which is in line with common practice in the literature.

2.2 Intangible vs. Traditional Value: Pricing Errors

This section examines the ability of the traditional and intangible value factors to price standard test portfolios. We begin by plotting the monthly returns to the intangible value ($\mathrm{HML}^{\mathrm{INT}}$) and traditional value ($\mathrm{HML}^{\mathrm{FF}}$) factors in Figure 2.1. As can be seen in the figure, the correlation between these two return series is high, with a full sample correlation coefficient estimate of 76.2%. We show that this correlation is high enough for intangible value to capture the "value effect," but low enough to allow intangible value to offer superior performance.

For our main asset pricing tests, we employ a two-step process. First, for each test asset i, we estimate betas from time-series regressions of portfolio excess returns on the risk factors

$$R_{it} = \alpha_i + \beta_{ik}\mathbf{k}_t + \epsilon_{it}, \tag{2.3}$$

where $\mathbf{k}_t$ is the vector of risk factors. These are MktRF, SMB, HML, and MOM for the three-factor model and the same factors plus RMW and CMA for the five-factor model.

Next, for each risk factor k, we estimate risk prices by running a cross-sectional regression of average excess returns on the estimated betas $\hat{\beta}_{ik}$

$$\mathbb{E}[R_{it}] = \eta_i + \hat{\beta}_{ik}\lambda_k + \nu_i. \tag{2.4}$$

The first two columns of Table 2.3 present the results for the Fama and French (1992, 1993) three-factor model plus momentum using the traditional value factor (column 1) and the intangible value factor (column 2). The test assets for these models are the standard size, book-to-market and momentum portfolios. As can be seen in the table, the intangible value factor reduces the alpha of this model by 5.4%, and reduces the root mean squared error by 3.7%. The χ^2 test rejects that the alphas from two models are different, and we conclude that intangible value prices standard test assets at least as well as traditional value in the three-factor model plus momentum.

Panels A and B of Figure 2.2 plot the results of these two models and report the mean absolute pricing errors, which HML^{INT} reduces by 2.2%. The figure shows that the fit of the two models is very similar for all test portfolios. One portfolio that has a smaller pricing error in the intangible model is S1B5, or Small Value. This portfolio displays high average returns. The higher loading on intangible value relative to traditional value brings the portfolio's predicted and actual returns closer in the intangible value model. Overall, despite putting HML^{INT} on unequal footing relative to HML^{FF} by requiring the book-to-market sorts to occur at the industry level (unlike the test assets), the models using within-industry-sorted HML^{INT} perform as well as or better than the models using traditional value.

The last two columns of Table 2.3 display the results for the Fama and French (2015) five-factor model plus momentum, which adds the conservative minus aggressive (CMA) investment factor and the robust minus weak (RMW) profitability factor. For this model, we also include the Fama and French investment and profitability portfolios as additional test assets. In the five-factor model with momentum, the coefficient on the traditional value factor is not statistically significant, while the intangible value factor retains significance at over the 1% level. Root mean squared errors are also smaller using HML^{INT}. The χ^2 test rejects that the alphas from the two models are different. Panels C and D of Figure 2.2 display the results visually, and report the smaller mean absolute pricing error for the intangible value model. We conclude that the intangible value factor does at least as well in pricing standard test assets as traditional value in both the classic three-factor model and the recently popularized five-factor model.

Figure 2.1 shows that there is a substantial commonality between the traditional and intangible value portfolios. To further draw out the unique pricing ability of intangible value, we additionally construct two distinct intangible value portfolios. The first, HML^{IME}, sorts firms only based on intangible assets relative to market equity. Table 2.4 presents the results for the three- and five-factor models plus momentum when this portfolio is used both in addition to the traditional value factor and on its own. The main message of this table is that an intangible-only value factor prices assets just as well as the traditional value factor.

The second decomposition we provide uses a portfolio, HML^{UINT}, which is long stocks that are *uniquely* in the long leg of HML^{INT} (that is, not in the long leg of HML^{FF}), and similarly goes short stocks which are in the short leg of HML^{INT} but either neutral or long in HML^{FF}. On average, about 20% of firms are used to construct HML^{UINT}, with about 60% coming from the long leg of intangible value, and 40% from the short leg. These fractions are all quite stable over time. As traditional value is not sorted within industries, we do not sort within industries when constructing the intangible value series used to construct HML^{UINT}. Table 2.5 presents the results for the three- and five-factor models plus momentum when this portfolio is used both in addition to the traditional value factor and on its own. χ^2 tests show that the difference in alphas from a three-factor model with traditional value and HML^{UINT} is statistically significant at the 1% level. We also find that the alphas in the three- and five-factor models with HML^{UINT} are larger than in the models with traditional value.

Our main results are produced with all industries in order to be as consistent as possible with the test assets and factor portfolios posted on the Fama and French data library – the series most widely utilized by researchers.[12] In the Online Appendix, we present our main results (including the analog of Table 2.3) without financials, utilities, and industries with SIC codes above 9,000. We show that intangible value also generates lower pricing errors using the smaller number of industries.

This section established that the intangible value factor prices standard test assets in the three- and five-factor models plus momentum with lower errors on average, and with alphas that are not significantly different, relative to the traditional value factor. This is true despite the fact that the 25 size and book-to-market test asset

[12] Several studies of the cross section of equity returns drop financials, utilities, and industries with SIC codes above 9,000. However, the Fama and French factors include all industries as noted in the authors' online documentation. We additionally verify that our replication of HML is substantially better when all industries are included.

portfolios are formed using the traditional book-to-market measure, and also that the intangible value factor sorts firms on total book (intangible plus recorded) to market equity within industries prior to value weighting each leg of the HML$^{\text{INT}}$ portfolio. When decomposing value into its traditional and intangible components using either HML$^{\text{IME}}$ or HML$^{\text{UINT}}$, we find that these more isolated intangible value portfolios alone produce similar pricing errors to traditional value. Tests for differences in alphas for the models with HML$^{\text{INT}}$ and HML$^{\text{IME}}$ as compared to the models with traditional value are indistinguishable. We conclude that intangible value appears to capture the value effect at least as well as or better than traditional value.

2.3 Intangible vs. Traditional Value: Performance

Figure 2.1 shows that the traditional and intangible value factors are highly correlated. The previous section documented that intangible value appears to capture the value effect at least as well as or better than traditional value. In this section, we show that there is enough independent variation in the two value factors to allow for substantial outperformance by the intangible value factor.

Table 2.6 documents the outperformance of intangible value relative to traditional value using single factor HML models. Panel A shows the results from a model of HML$^{\text{INT}}$ regressed on the HML$^{\text{FF}}$ factor. We present results for the full sample and for subsamples covering the pre-internet era from 1975 to 1994, the internet era pre-crisis from 1995 to 2006, and the crisis and post-crisis era from 2007 to 2018. The alpha of HML$^{\text{INT}}$ over HML$^{\text{FF}}$ is 3.86% in the full sample and statistically significant at the 1% level. This outperformance is sizable given the apparent close relationship between the two factors. However, this fact is also reasonable as the appraisal ratio (α/RMSE) is 0.91. Interestingly, the alpha is fairly stable over time, and is statistically significant in all subsamples, though at a somewhat lower level in the most recent subsample.

Turning to Panel B, which shows the results for the converse model in which HML^{FF} is regressed on the HML^{INT} factor, we see that the alpha is -3.03% and statistically significant at the 1% level for the full sample. Looking at the subsamples, the third and fourth columns show that the most prominent underperformance of HML^{FF} relative to HML^{INT} comes in the recent periods of 1995 to 2006 and 2007 to 2018. The recent underperformance is notable because the post-crisis era has been one of the worst periods for the traditional value strategy. We find that the intangible strategy performed significantly better from 2007 to 2018, by 3.59%.[13]

Next, we compare the outperformance of our measure of intangible value over an intangible value factor constructed using the Peters and Taylor (2017) method employed by Park (Forthcoming). To construct this alternative intangible value factor, $\text{HML}^{\text{PTINT}}$, we sort firms unconditionally across all industries and accumulate 30% of (SG&A-R&D) plus 100% of R&D. Table 2.7 presents the results. Our HML^{INT} factor has a positive alpha of 2.42% over $\text{HML}^{\text{PTINT}}$ in the full sample. The alpha is positive in all subsamples, though not statistically significant in the post-crisis era. The alphas of $\text{HML}^{\text{PTINT}}$ with respect to our intangible value factor are all negative, but largely not significant. We conclude that our intangible value factor outperforms the factor used in Park (Forthcoming).

Table 2.8 examines the outperformance of the two decompositions of intangible value, HML^{IME} and HML^{UINT}. As expected, the two portfolios that isolate the effect of intangibles display more independent variation from traditional value, implied by the lower R^2 compared to corresponding columns in Table 2.6. The full sample alpha is larger for both HML^{IME} (4.95%) and HML^{UINT} (4.71%). Moreover, the alphas for these factors are larger than alphas from the baseline intangible value regression and are also statistically significant in the post financial crisis period. Similar to the case

[13] The Online Appendix contains results using a traditional value factor that is sorted within industries and finds the same patterns, with slightly smaller magnitudes for outperformance as expected.

of the baseline intangible value portfolio, the outperformance of portfolios that isolate the effect of intangibles appears to be strongest in the pre-crisis internet era from 1995 to 2006.

Eisfeldt and Papanikolaou (2013b) showed that firms with more organization capital to physical capital earned positive excess returns even when controlling for the Fama and French three factors plus momentum. They also use accumulated SG&A to measure the stock of intangible organization capital. However, that factor is substantially different from intangible value, which is not surprising given that the organization capital factor compares two book values, while our intangible value factor compares book value (including intangibles) to market value. Table 2.9 clearly shows that the intangible value factor is quite different from the organization capital factor. In the full sample, the R^2 in a regression of intangible value on the organization capital factor from Eisfeldt and Papanikolaou (2013b) is negligible (0.09%). We conclude that although both factors provide evidence of the importance of intangibles for asset pricing, they capture different effects both conceptually and empirically.

Table 2.10 displays performance statistics for various value factors: $\mathrm{HML}^{\mathrm{FF}}$, $\mathrm{HML}^{\mathrm{INT}}$, $\mathrm{HML}^{\mathrm{IME}}$, a portfolio that is long $\mathrm{HML}^{\mathrm{INT}}$ and short $\mathrm{HML}^{\mathrm{FF}}$, and a portfolio that is long $\mathrm{HML}^{\mathrm{IME}}$ and short $\mathrm{HML}^{\mathrm{FF}}$. We show results for average returns, volatility, confidence intervals, and Sharpe ratios. For the long-short portfolios, we add information and appraisal ratios using intangible value as the traditional value benchmark and vice versa for traditional value. The top panel shows that the traditional value factor had a positive and statistically significant return over the full sample. However, the significance is mainly driven by the earliest two subsamples of 1975 to 1994 and 1995 to 2006. In fact, the average returns to $\mathrm{HML}^{\mathrm{FF}}$ are (not significantly) negative in the most recent subsample (2007 to 2018). In contrast, the average returns to intangible value are substantially larger in magnitude and significance over the full sample, with the positive returns exhibiting higher significance through 2006. In the most recent subsample, average returns are positive but not statistically significant. We find that HML$^{\text{INT}}$ still significantly outperforms HML$^{\text{FF}}$, and as shown in Table 2.6, this outperformance actually increases in recent years. HML$^{\text{IME}}$ exhibits even higher returns and lower volatility across all periods, resulting in the highest Sharpe ratio.

The second to last panel displays portfolio performance statistics for the long intangible value, short traditional value strategy. This strategy has a positive and statistically significant average return over the full sample (2.11%), and a Sharpe ratio of 0.32. Moreover, the returns performance of this strategy has been improving over time, and most of the significantly positive outperformance actually comes from the most recent subsample when traditional value underperformed. During the 2007 to 2018 subsample, the Sharpe ratio of the long-short strategy is 0.62. The appraisal ratio, which compares the performance of HML$^{\text{INT}}$ and HML$^{\text{FF}}$, is also positive throughout the entire sample, indicating HML$^{\text{INT}}$'s superior performance. The bottom panel examines the performance of a portfolio that is long HML$^{\text{IME}}$ and short HML$^{\text{FF}}$. The return of this portfolio is significantly positive at 2.86% over the full sample, which is again mainly driven by the substantial outperformance of HML$^{\text{IME}}$ over HML$^{\text{FF}}$ in the most recent subsample. The average return of this long-short strategy is 5.05% in the most recent subsample with a Sharpe ratio of 0.70. Consistent with this, the appraisal ratio between HML$^{\text{IME}}$ and HML$^{\text{FF}}$ is positive throughout all periods.

Figure 2.3 plots the cumulative returns for several long-short strategies for the full sample and for the subsamples starting in 1995 (post internet era) and 2007 (post Great Financial Crisis). The top panel plots the cumulative returns to investing one dollar in either HML$^{\text{FF}}$ or HML$^{\text{INT}}$, and clearly shows the superior returns to HML$^{\text{INT}}$ in the full sample and in each subsample. The middle panel plots the cumulative returns to the portfolio that is long HML$^{\text{INT}}$ and short HML$^{\text{FF}}$. Again, the outperformance of HML$^{\text{INT}}$ is apparent. In terms of the subsamples, it appears that the post-internet era is an important driver of the outperformance, as is the post-crisis era during which social media firms thrived. This is consistent with the growth of intangible capital documented in prior studies.

The bottom panel shows the cumulative returns to the intangible and traditional value strategies, along with the cumulative returns to the factors from the three- and five-factor models plus momentum for comparison. Over the full sample, the intangible value factor's performance is of a very similar magnitude to the best performing factor, momentum (UMD), while exhibiting much lower volatility (and no extreme draw-downs as observed in the momentum crash of 2007). Intangible value's performance is clearly far superior to any other factor in the Fama and French (2015) five-factor model. Between 1995 and 2018, the intangible value factor displays the highest performance of any of the long-short portfolios. In the most recent subsample, intangible value outperformed all other factors with the exception of the profitability factor (RMW).

Figure 2.4 decomposes the outperformance of intangible value into the contributions of the superior long leg and the superior short leg by plotting cumulative returns to the differences in each value portfolio's long and short legs, respectively. We present long and short leg returns for the full sample as well as for the post-internet subsample and the post-crisis subsample. We find that going long the short leg of traditional value and short the long leg of traditional value appears to be a fairly low volatility, positive return strategy. This implies that intangible value avoids shorting firms with book anchors that understate total book capital by not incorporating intangibles.

Table 2.11 displays alphas of the traditional and intangible value factors in the three- and five-factor models plus momentum. We include results for the baseline intangible value factor and for the two factors that isolate the effect of intangible capital. In the three-factor model plus momentum, the alpha for traditional value is negative but not significant. In contrast, the alpha for HML^{INT} is 2.92%, and is highly statistically significant at the 1% level. The alpha for HML^{IME}, which sorts firms using the ratio of intangible assets to market equity, is 3.87% and significant at the 1% level. The alpha for HML^{UINT}, which only contains stocks unique to the HML^{INT} long or short leg, is positive but not significant.

In the five-factor model plus momentum, the alpha for traditional value is negative whereas the alphas for the intangible value factors except HML^{UINT} are positive and strongly significant. This is notable as Fama and French (2015) find that the original value factor becomes redundant when the investment and profitability factors are added, although, as shown in Table 2.3, this is not true for HML^{INT}. The intangible value factor has a positive and significant loading on RMW, or the robust minus weak factor, meaning that the intangible value factor comoves with the returns to firms with stronger profitability. This is consistent with the evidence we present in the next section that the long leg of the intangible value factor, unlike the traditional value factor, tends to contain more productive firms, and vice versa for the short leg. We conclude from Table 2.11 that the intangible value factors all have positive and significant alphas in the three- and five-factor models plus momentum, with the exception of HML^{UINT}, for which the positive alphas are not significant.

2.4 How do Intangible and Traditional Value Differ?

Intangible value generates similar pricing errors relative to traditional value but outperforms significantly, leading to a large Sharpe ratio for a strategy that is long intangible value and short traditional value. In this section, we investigate the properties of firms that are in the long and short legs of intangible, vs. traditional, value. Table 2.12 presents results on characteristics of firms that are in the short, neutral, and long legs of intangible value and traditional value. Here, we report the time-series average of the median firm characteristic within each bucket. Not surprisingly, the first two rows show that there are larger differences in total book to market equity for intangible value, and larger differences in recorded book to market equity for traditional value, across the three possible portfolio rankings. Intangible value tends to be long slightly smaller firms, and short slightly larger firms than traditional value. This is consistent with their loadings on SMB in the three- and five-factor models, which are positive for intangible value and negative for traditional value. Importantly, intangible value has a positive and significant alpha of 2.92% controlling for the market, size, value, and momentum, as shown in Table 2.11. Row four of Table 2.12 shows that the expected pattern for intangible capital to book assets holds for the intangible value portfolio legs. On average, firms with higher intangible capital to recorded book assets appear in the long leg, and firms with a lower ratio of intangible capital to recorded book assets appear in the short leg. We observe the opposite pattern for the traditional value portfolio; the long leg has lower intangible capital to recorded book assets than the short leg. Row five shows that a similar pattern holds for intangible capital to sales, which is intuitive because intangible capital is measured as accumulated SG&A expenses.

Rows six and seven in Table 2.12 document that productivity tends to be increasing in $\mathrm{B/M}^{\mathrm{INT}}$, and decreasing in $\mathrm{B/M}^{\mathrm{FF}}$. Thus, $\mathrm{HML}^{\mathrm{INT}}$ is long higher productivity firms and short lower productivity firms, while $\mathrm{HML}^{\mathrm{FF}}$ is long lower productivity firms and short higher productivity firms. Productivity, measured as sales to recorded assets, is monotonically increasing across the intangible value legs, and monotonically decreasing across the traditional value legs. Using Solow (1955, 1957) residuals to measure productivity yields slightly more mixed results, but still favors intangible value. The residuals are fairly flat across the intangible portfolio legs. However, the Solow residuals are monotonically decreasing across the traditional value legs, meaning that traditional value is short firms with higher Solow residuals and long firms with lower Solow residuals. Row eight shows that $\mathrm{HML}^{\mathrm{INT}}$ is long firms with higher sales to stockholder's equity and short firms with lower sales to stockholder's equity, while HML^{FF} displays the opposite pattern.

In terms of alternative valuation measures, row nine shows that intangible value is long firms with slightly lower price to diluted earnings (P/E ratios) excluding extraordinary items relative to traditional value, and short firms with higher P/E ratios. Row ten shows that the two portfolios have similar patterns for price to sales. We conclude that including intangible capital aligns the B/M measure of value with measures that use P/E.

Rows eleven and twelve focus on measures of financial soundness. While intangible and traditional value have fairly similar patterns of debt to book assets across their long and short legs, traditional value tends to be long firms with much higher debt to EBITDA, indicating that firms in the long leg of traditional value may be less financially sound. Row thirteen shows that the dividend yield increases across terciles for both intangible and traditional value, with a slightly steeper slope for HML^{FF}.

Next, we report statistics related to the investment (CMA) and profitability (RMW) factors in the five-factor model plus momentum. Row fourteen shows that both intangible and traditional value tend to be long firms with lower investment to physical capital (capital expenditures to PP&E), consistent with the arguments in Hou et al. (2015). Row sixteen shows that intangible value, unlike traditional value, tends to be long firms with higher gross profit to total assets, and short firms with lower gross profit to book assets. Instead, traditional value tends to be short more profitable firms and long less profitable firms by this measure. This is consistent with the evidence in Table 2.11 that intangible value, unlike traditional value, loads positively on the RMW factor.

Our study is aimed at documenting the pricing ability and performance statistics of an intangible value factor that is constructed efficiently and minimizes biases due to accounting differences across industries. We largely leave the underlying economic reasons for intangible value's outperformance over traditional value to future work. One reason behind intangible value's outperformance might be behavioral.[14] Value firms may be underpriced, and intangible assets may be more sensitive to mispricing. Another explanation is that intangible value better captures firms' exposure to technology shocks that displace the value of assets in place including intangible assets, but increase the value of growth opportunities.[15] This is consistent with the results in Goncalves and Leonard (2020), which finds that including intangible capital improves the ability of book equity to capture fundamental equity values by 30% in recent data. We also find some support for the latter explanation by comparing the exposures of HML^{INT} vs. HML^{FF} to technology shocks. We measure technology shocks following Kogan et al. (2020) using the market value of patents. The last row of Table 2.12 reports loadings on the these shocks, controlling for market returns. The spread in technology risk exposures between the bottom and the top 30% of firms increases by 42% (from 0.12 to 0.17) when intangible assets are included.[16]

In summary, the analysis of firm characteristics across book-to-market terciles for intangible and traditional value seems to indicate why intangible value may outperform traditional value. "Value traps" are value firms with high book to market ratios whose market values do not recover. As the fundamentals (measured by productivity and alternative valuation ratios) seem better for the long leg of intangible value (and worse for the short leg), relative to traditional value, it may be that intangible value avoids these value traps. For instance, Finish Line was sorted uniquely in the long leg of traditional value for 30% of the period it was traded. While the stock appeared cheap using traditional B/M, it suffered from lagging performance behind competitors

[14] See Daniel and Titman (1997) and Golubov and Konstantinidi (2019).

[15] See, for example, Papanikolaou (2011); Kogan and Papanikolaou (2014); Kogan et al. (2020).

[16] Note that the difference in exposures between the long and short legs of value are statistically significant at the 10% (traditional value) and 1% level (intangible value). The difference in the spread in exposures, however, is not significant, perhaps due to the fact that the data for this exercise are annual and aggregate.

(including online retail) and never recovered until its acquisition in 2018. Similarly, by including investment in intangible assets, intangible value may outperform traditional value by avoiding short positions in firms whose book values do not accurately anchor their fundamental value. Well-known companies such as Target, Nordstrom, and Estee Lauder have consistently been sorted into the short leg of traditional value despite consistently investing in systems and customer-related intangibles. In most periods, intangible value in fact takes a long position in these stocks, amplifying the difference in returns between the two value factors.

It is also interesting to examine how persistent the differences in positions between HML^{INT} and HML^{FF} are. Table 2.13 addresses this question by reporting the empirical transition matrices and the respective stationary distributions showing the probability that a firm is uniquely in a particular leg of either the intangible or traditional value portfolio. The first matrix shows transition probabilities for firms that are uniquely in the long leg of intangible value. Such firms are in the top 30% of firms ranked by B/M^{INT}, but in the bottom 70% of firms ranked by recorded B/M. These unique positions are fairly persistent; with 58% probability, a firm in the long leg of intangible value that is either neutral or short in traditional value remains uniquely in the long leg of intangible value in the following period. This implies that differences between HML^{INT} and HML^{FF} are driven in part by persistent differences in the rankings of firms. The remaining three matrices show that the probability of remaining uniquely in the short leg of HML^{INT}, or uniquely in the long or short leg of HML^{FF}, are all over 50%. Note that the actual persistence of positions that would be used to infer turnover costs are much higher as Table 2.13 considers only the persistence of the positions that drive the return differences between intangible and traditional value. The implied stationary distributions show that firms spend between 7% and 16% of the time in positions that differ between intangible and traditional value.

2.5 Conclusion

The traditional value investing strategy, which relies on using firms' book assets as the fundamental anchor of value, has lost its edge in recent years. This trend may be due to the increasing importance of intangible capital, which is not incorporated into the traditional measure of book assets. We show that a value portfolio that adds intangible capital to book assets prior to sorting provides much stronger performance in all periods. The intangible value factor also prices standard test assets with similar pricing errors as the traditional value factor.

We emphasize sorting firms within industries when constructing intangible value because industry standards for allocating costs vary across industries. Similarly, we advocate doing the within-industry sort based on an intangible capital stock that is formed using 100% of SG&A, as opposed to 30% of SG&A and 100% of R&D, due to the large differences in R&D reporting practice across industries and the resulting bias that can arise from replacing missing R&D with zeros. Using 100% of SG&A also better accounts for organization, brand and customer capital, the importance of which can be substantial in many industries.

We also find that long-short strategies that better isolate the effects of intangible capital on value continue to price standard test assets and yield positive and significant alphas. Lastly, we document that, on average, intangible value is long firms with better fundamentals (productivity, earnings, and profitability) relative to traditional value.

Taken together, our findings show that asset pricing researchers should consider adjusting the value factor and accompanying test assets to incorporate intangible capital. Practitioners can also use the intangible value factor to implement a highly profitable relative value strategy that is long intangible value and short traditional value. This strategy has exhibited strongly positive returns and a high Sharpe ratio, especially in recent years when traditional value has underperformed.

2.A Figures and Tables

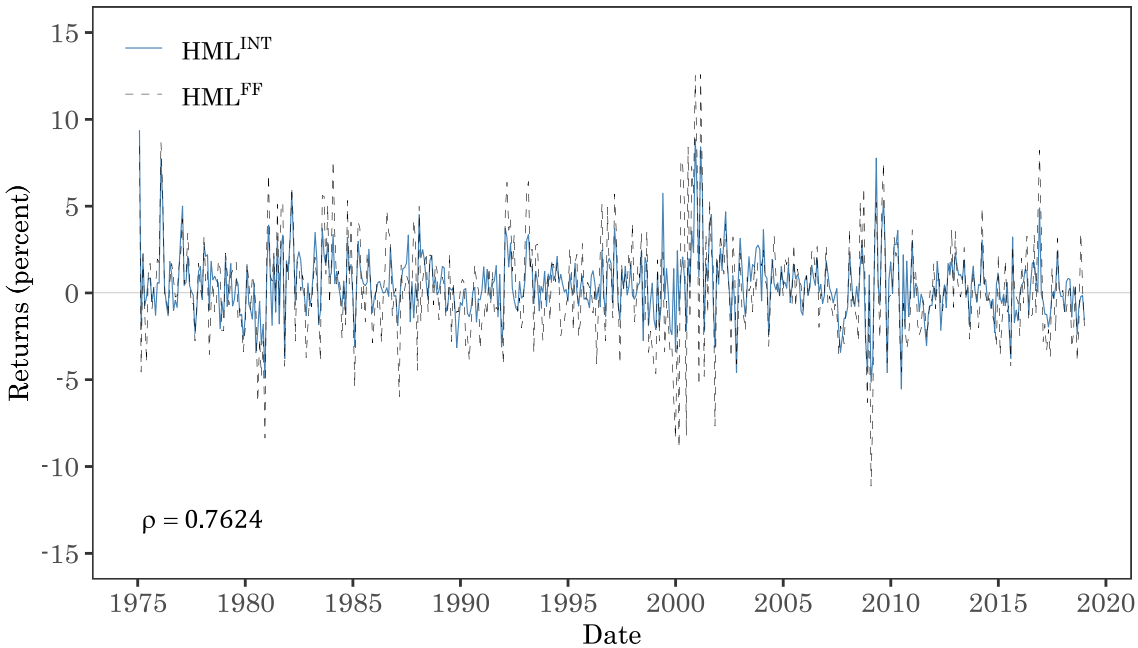


Figure 2.1: Relationship between Intangible and Traditional Value

Description: This figure plots monthly returns for HML^{FF} and HML^{INT} from 1975 to 2018. The HML^{FF} portfolio mimics the risk factor in returns related to book-to-market equity, and is calculated as the difference between the returns on high-B/M portfolios and the returns on low-B/M portfolios. HML^{FF} returns are downloaded from Ken French's website. HML^{INT} adds intangible assets to the book equity term of the book-to-market equity ratio and conduct portfolio sorts within industries. Further details on factor construction can be found in Section 2.1 and the Online Appendix. ρ reports the correlation between the two returns for the full sample period.
Interpretation: The full sample correlation coefficient between traditional and intangible value is 76.2%. This correlation is high enough for intangible value to capture the value effect but low enough to allow intangible value to offer superior performance.

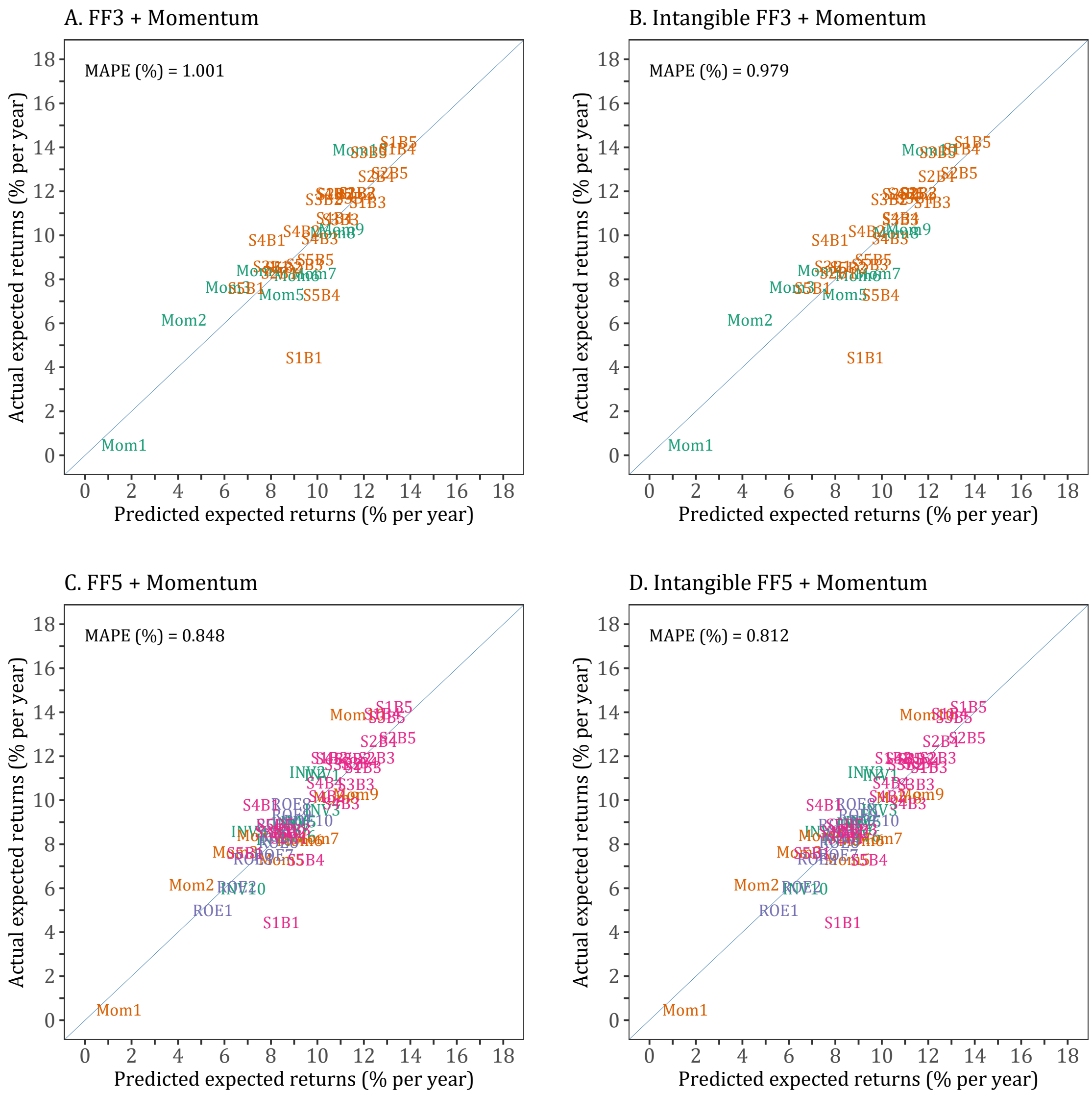


Figure 2.2: Cross-sectional Asset Pricing Tests – Intangible and Traditional Value

Description: This figure shows the cross-sectional asset pricing tests from the Fama and French (1992, 1993, 2015) three-factor and five-factor models. Panel A plots realized mean excess returns of 25 size and book-to-market-sorted portfolios and 10 momentum portfolios against the mean excess returns predicted by the FF3 + momentum model. Panel C plots realized mean excess returns of 25 size and book-to-market sorted portfolios, 10 momentum portfolios, 10 portfolios sorted on operating profitability, and 10 portfolios sorted on investment, against the mean excess returns predicted by the FF5 + momentum model. Panels B and D replace HML^{FF} with HML^{INT}. The sample is monthly from 1975 to 2018. Returns are reported in percent per year.
Interpretation: Replacing traditional value with intangible value reduces mean absolute pricing errors, showing that intangible value does at least as well in pricing standard test assets as traditional value.

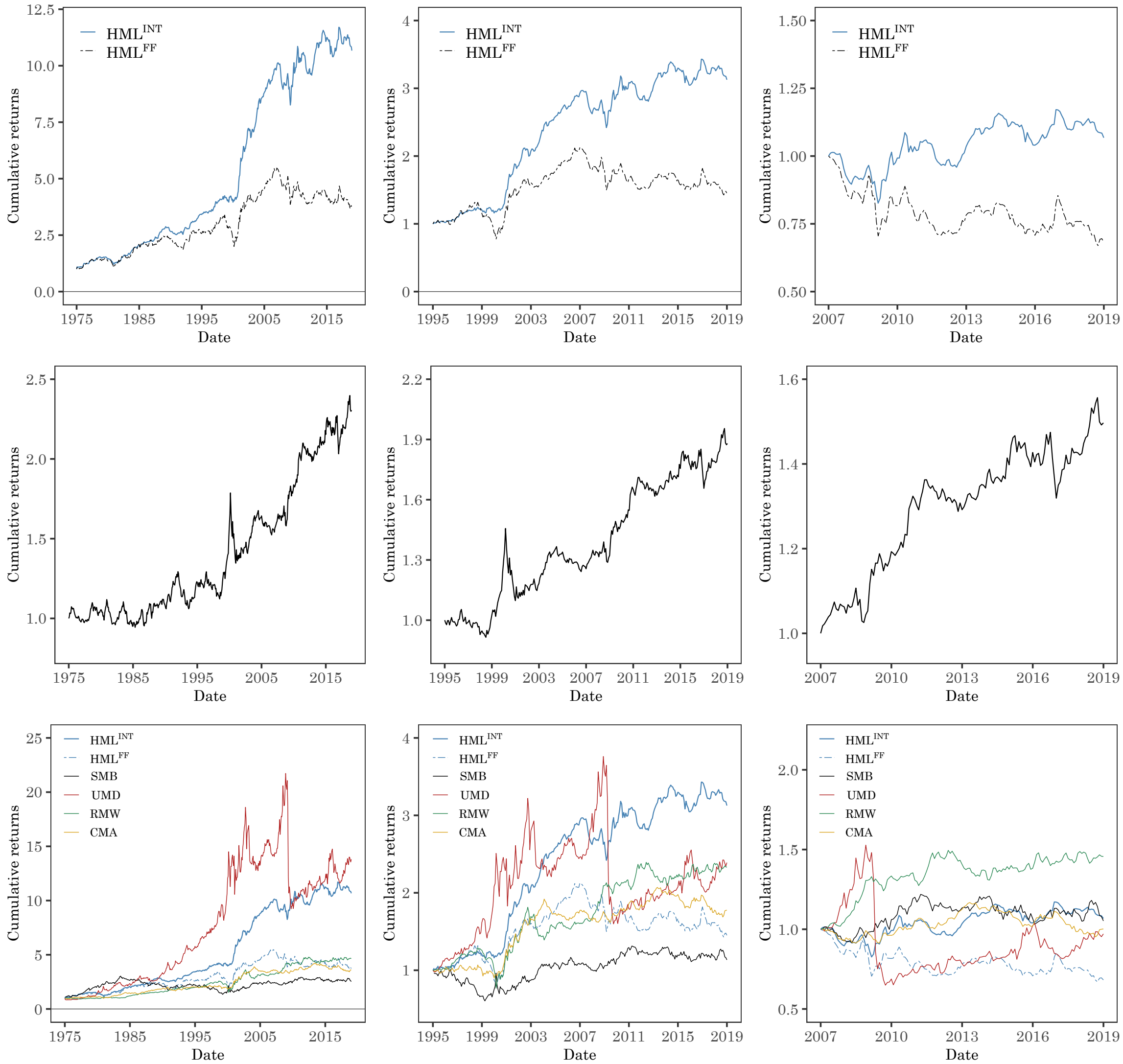


Figure 2.3: Performance of Intangible Value

Description: The top panel plots the cumulative returns of one dollar invested in the HML^{FF} and HML^{INT} portfolios from the beginning of 1975, 1995, and 2007. The middle panel plots the cumulative returns of one dollar invested in the portfolio that is long the HML^{INT} portfolio and short the HML^{FF} portfolio. The bottom panel plots the cumulative returns of one dollar invested in HML^{INT}, the Fama and French five factors, and momentum.

Interpretation: Intangible value outperforms traditional value in both the full sample period and recent sub-samples. A long-short portfolio of intangible and traditional value also has positive returns. Lastly, intangible value exhibits similar performance as the top-performing momentum factor without suffering from as large a drawdown.

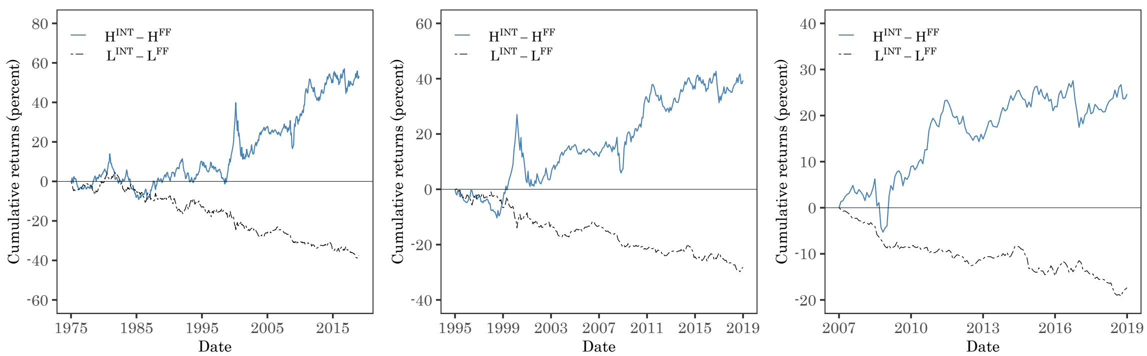


Figure 2.4: Decomposing the Outperformance of Intangible Value

Description: This figure plots the cumulative returns of a portfolio that is long the long leg of HML^{INT} and short the long leg of HML^{FF} (solid line), as well as the returns of a portfolio that is long the short leg of HML^{INT} and short the short leg of HML^{FF} (dashed line). Each panel plots percent returns from the beginning of 1975, 1995, and 2007.

Interpretation: The long leg of intangible value outperforms the short leg of traditional value, while the short leg of intangible value underperforms the short leg of traditional value. This implies that a portfolio that shorts the traditional value factor appears to be a fairly low volatility, positive return strategy.

		Market-wide ($\gamma_{B,t}$)	Across-industry ($\gamma_{1,t}$)	Within-industry ($\gamma_{2,t}$)
log(B/M)	(i)	0.38		
		(5.82)		
	(ii)		-0.32	0.44
			(-1.20)	(8.65)
log(B^{INT}/M)	(i)	0.40		
		(7.56)		
	(ii)		0.04	0.43
			(0.24)	(8.75)

Table 2.1: Value and the Cross Section of Stock Returns

Description: This table reports average slopes and T-statistics of monthly single variable cross-sectional regressions following Asness et al. (2000). $r_{i,t}$ refers to monthly stock returns of firm i at time t. $X_{i,t}$ is log(B/M) or log(B^{INT}/M) for firm i at time t, while $X_{I,i,t}$ is the average log(B/M) or log(B^{INT}/M) for the industry I of firm i. B/M is formed each July using prior end of December's market equity and prior year's BE. Industry definitions are adopted from the Fama-French 12 industry classifications. The sample is from January 1975 to December 2018.

Market-wide regression equation: $$r_{it} = \gamma_{A,t} + \gamma_{B,t}X_{i,t} + \epsilon_{i,t} \quad \text{(i)}$$

Industry regression equation: $$r_{it} = \gamma_{0,t} + \gamma_{1,t}X_{I,i,t} + \gamma_{2,t}(X_{i,t} - X_{I,i,t}) \quad \text{(ii)}$$

Interpretation: Book-to-market's ability to predict stock returns is almost entirely driven by within-industry variation. Measuring value within industries thus reduces noise and exposure to unpriced risk, increasing Sharpe ratios.

	Panel A					Panel B				
	xrd/sale					*xsga/(cogs+xsga)*				
						Full Sample			1995-2006	2007-2018
Industry	% NA	% 0	Mean	s.d.	p50	Mean	s.d.	p50	Mean	Mean
Consumer Nondurables	68.29	3.42	3.15	21.61	0.78	27.31	15.31	25.50	30.77	31.65
Consumer Durables	34.07	1.27	9.12	206.87	1.79	23.36	14.82	20.29	24.55	25.82
Manufacturing	34.86	1.34	10.70	299.35	1.60	21.92	13.33	19.42	23.18	22.24
Energy	80.77	6.93	43.88	1114.35	0.28	28.99	24.21	21.61	26.12	24.24
Chemicals	22.32	0.56	9.09	112.04	2.18	29.92	18.79	25.49	30.43	26.59
Business Equipment	11.74	0.98	71.20	3704.26	9.62	44.10	21.72	41.06	48.49	47.80
Telecommunications	79.26	2.93	95.33	1522.05	2.22	40.11	19.15	39.49	40.43	40.76
Utilities	99.34	0.34	0.08	0.18	0.00	20.02	17.42	16.65	19.21	33.49
Wholesale and Retail	39.29	53.61	1.05	43.56	0.00	24.42	14.95	22.47	24.62	23.65
Healthcare	12.38	5.87	1511.40	29821.61	11.73	53.67	25.64	51.64	53.35	62.65
Finance	95.36	1.91	17.32	109.16	0.38	40.50	20.76	36.34	38.15	52.96
Other	77.09	8.56	77.36	1203.57	0.48	27.09	21.58	20.55	27.13	28.07

Table 2.2: Descriptive Statistics

Panel C

Industry	N	% Mkt Cap	B/M					B^{INT}/M				
			Mean	s.d.	p10	p50	p90	Mean	s.d.	p10	p50	p90
Consumer Nondurables	11843	7.24	1.10	1.19	0.24	0.79	2.23	4.62	8.32	0.52	2.21	10.63
Consumer Durables	5181	2.67	0.99	0.97	0.25	0.73	2.02	3.66	6.01	0.50	1.85	8.52
Manufacturing	23387	10.27	1.04	1.05	0.27	0.78	2.00	3.32	4.75	0.49	1.84	7.50
Energy	7874	9.04	0.88	0.89	0.22	0.68	1.65	1.48	1.91	0.34	0.97	2.97
Chemicals	4489	4.39	0.78	0.77	0.20	0.56	1.57	2.98	4.80	0.40	1.44	6.73
Business Equipment	29404	14.66	0.67	0.66	0.14	0.49	1.35	2.29	3.67	0.30	1.21	5.05
Telecommunications	3418	6.35	0.84	1.12	0.14	0.57	1.66	1.63	2.82	0.23	0.86	3.23
Utilities	6468	5.43	1.12	0.54	0.61	1.03	1.73	1.18	1.05	0.61	1.04	1.80
Wholesale and Retail	19981	8.40	1.04	1.14	0.23	0.72	2.12	6.32	13.22	0.53	2.32	14.97
Healthcare	16034	8.90	0.50	0.62	0.09	0.33	1.03	1.35	2.46	0.17	0.62	3.09
Finance	30239	14.97	1.02	1.13	0.36	0.81	1.75	1.86	4.78	0.43	1.08	3.39
Other	22139	7.68	0.98	1.50	0.20	0.68	1.89	2.54	5.28	0.30	1.19	5.43

Table 2.2: *Continued*

Description: This table summarizes key variables related to the calculation of intangible capital. *xrd*, *sale*, *cogs*, and *xsga* are Compustat variables for R&D expenditures, sales, SG&A expenditures, and cost of goods sold. % NA and % 0 refer to the fraction of firms reporting missing numbers or zero, respectively. In Panel C, N is the total number of firm-year observations for each industry, and % Mkt Cap is the fraction of each industry's market capitalization for the full period. B/M is the traditional book-to-market ratio and B^{INT}/M denotes the intangible-adjusted book-to-market ratio used to construct HML^{INT}. We report statistics using annual data at time of portfolio formation (June of each year) from 1975 to 2018. All fractions except B/M and B^{INT}/M are denoted in percent.

Interpretation: Panels A and B show industry-level differences in practices for reporting R&D expenses, providing support for our methodology of accumulating total SG&A expenditures, which in Compustat includes R&D even when it is separately reported. Panel C shows how the impact of incorporating intangibles into book to market ratios varies across industries.

	(1)	(2)	(3)	(4)
α (%)	13.28	12.56	8.59	9.12
	(4.15)	(3.94)	(2.89)	(3.06)
β_{MktRF}	-0.38	-0.33	-0.04	-0.08
	(-1.18)	(-1.02)	(-0.12)	(-0.26)
β_{SMB}	0.18	0.19	0.24	0.23
	(1.36)	(1.38)	(1.78)	(1.75)
$\beta_{HML^{\mathrm{FF}}}$	0.30		0.24	
	(2.35)		(1.92)	
$\beta_{HML^{\mathrm{INT}}}$		0.29		0.30
		(2.87)		(2.88)
β_{UMD}	0.54	0.55	0.53	0.54
	(2.79)	(2.80)	(2.74)	(2.77)
β_{RMW}			0.32	0.32
			(2.87)	(2.90)
β_{CMA}			0.18	0.16
			(1.95)	(1.69)
Adj. R^2	73.14	75.12	78.74	79.84
RMSE	0.43	0.42	0.34	0.33
Prob > χ^2		0.19		0.24

Table 2.3: Pricing Errors – Intangible Value vs. Traditional Value

Description: This table represents pricing results for the Fama and French (1992, 1993, 2015) three- and five-factor models plus momentum. In terms of test assets, columns (1) and (2) use 25 portfolios double-sorted on size and book-to-market and 10 portfolios sorted on momentum. Columns (3) and (4) additionally include 10 investment and 10 profitability portfolios. Fama and MacBeth (1973) T-statistics are reported in parentheses. Prob > χ^2 is the p-value of the test that the alpha from the model using $\mathrm{HML}^{\mathrm{INT}}$ is significantly different from the alpha from the model using $\mathrm{HML}^{\mathrm{FF}}$. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve.
Interpretation: Intangible value prices standard test assets at least as well as traditional value in the three and five-factor models plus momentum.

	(1)	(2)	(3)	(4)	(5)	(6)
α (%)	13.28	13.30	12.80	8.59	8.55	8.40
	(4.15)	(4.02)	(3.90)	(2.89)	(2.89)	(2.81)
β_{MktRF}	-0.38	-0.38	-0.34	-0.04	-0.03	-0.02
	(-1.18)	(-1.17)	(-1.04)	(-0.12)	(-0.11)	(-0.07)
β_{SMB}	0.18	0.18	0.18	0.24	0.23	0.23
	(1.36)	(1.36)	(1.35)	(1.78)	(1.77)	(1.77)
$\beta_{HML^{\mathrm{FF}}}$	0.30	0.30		0.24	0.25	
	(2.35)	(2.35)		(1.92)	(1.92)	
$\beta_{HML^{\mathrm{IME}}}$		0.16	0.31		0.22	0.27
		(1.01)	(2.73)		(1.46)	(2.34)
β_{UMD}	0.54	0.54	0.54	0.53	0.53	0.53
	(2.79)	(2.79)	(2.78)	(2.74)	(2.74)	(2.73)
β_{RMW}				0.32	0.32	0.33
				(2.87)	(2.93)	(2.95)
β_{CMA}				0.18	0.18	0.19
				(1.95)	(1.95)	(2.02)
Adj. R^2	73.14	72.22	72.56	78.74	78.32	78.67
RMSE	0.43	0.44	0.44	0.34	0.34	0.34
Prob $> \chi^2$		0.98	0.43		0.83	0.51

Table 2.4: Pricing Errors – Intangible Assets to Market Equity

Description: This table represents pricing results for the Fama and French (1992, 1993, 2015) three- and five-factor models plus momentum. In terms of test assets, columns (1) through (3) use 25 portfolios double-sorted on size and book-to-market and 10 portfolios sorted on momentum. Columns (4) through (6) additionally include 10 investment and 10 profitability portfolios. $\mathrm{HML}^{\mathrm{IME}}$ is the HML factor that replaces book-to-market with intangibles-to-market as the sorting variable. Fama and MacBeth (1973) T-statistics are reported in parentheses. Prob $> \chi^2$ is the p-value of the test that the alpha from the model using $\mathrm{HML}^{\mathrm{IME}}$ is significantly different from the alpha from the model using $\mathrm{HML}^{\mathrm{FF}}$. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve).
Interpretation: A value factor that sorts only on intangible assets to market equity prices assets just as well as the traditional value factor.

	(1)	(2)	(3)	(4)	(5)	(6)
α (%)	13.28	12.81	20.34	8.59	9.40	8.77
	(4.15)	(4.00)	(5.19)	(2.89)	(3.15)	(2.92)
β_{MktRF}	-0.38	-0.32	-0.89	-0.04	-0.09	-0.03
	(-1.18)	(-0.99)	(-2.47)	(-0.12)	(-0.28)	(-0.11)
β_{SMB}	0.18	0.17	0.16	0.24	0.22	0.23
	(1.36)	(1.29)	(1.20)	(1.78)	(1.68)	(1.74)
$\beta_{HML^{\mathrm{FF}}}$	0.30	0.27		0.24	0.25	
	(2.35)	(2.13)		(1.92)	(2.00)	
$\beta_{HML^{\mathrm{UINT}}}$		1.20	1.53		1.09	1.07
		(4.34)	(4.88)		(4.26)	(4.23)
β_{UMD}	0.54	0.52	0.44	0.53	0.51	0.49
	(2.79)	(2.64)	(2.25)	(2.74)	(2.59)	(2.51)
β_{RMW}				0.32	0.34	0.34
				(2.87)	(3.05)	(3.04)
β_{CMA}				0.18	0.20	0.27
				(1.95)	(2.14)	(2.72)
Adj. R^2	73.14	79.71	71.56	78.74	83.60	82.64
RMSE	0.43	0.37	0.44	0.34	0.30	0.31
Prob > χ^2		0.78	0.00		0.48	0.89

Table 2.5: Pricing Errors – Intangible Value with Unique Sort

Description: This table represents pricing results for the Fama and French (1992, 1993, 2015) three- and five-factor models plus momentum. In terms of test assets, columns (1) through (3) use 25 portfolios double-sorted on size and book-to-market and 10 portfolios sorted on momentum. Columns (4) through (6) additionally include 10 investment and 10 profitability portfolios. $\mathrm{HML}^{\mathrm{UINT}}$ is a factor that goes long firms that are in the long leg of $\mathrm{HML}^{\mathrm{INT}}$ but not in the long leg of $\mathrm{HML}^{\mathrm{FF}}$, and vice versa for the short leg ("unique" intangible factor). Fama and MacBeth (1973) T-statistics are reported in parentheses. Prob > χ^2 is the p-value of the test that the alpha from the model using $\mathrm{HML}^{\mathrm{UINT}}$ is significantly different from the alpha from the model using $\mathrm{HML}^{\mathrm{FF}}$. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve).
Interpretation: The alpha in the three- and five-factor models with $\mathrm{HML}^{\mathrm{UINT}}$ is larger than the models with traditional value, and the difference is significant at the 10% level for the three-factor model.

	Full sample	1975-1994	1995-2006	2007-2018
	(1)	(2)	(3)	(4)
A. $\mathbf{HML_t^{INT}} = \boldsymbol{\alpha} + \boldsymbol{\beta_{HML^{FF}}} \cdot \mathbf{HML_t^{FF}} + \boldsymbol{\epsilon_t}$				
α (%)	3.86	3.68	6.18	2.32
	(6.10)	(4.18)	(4.82)	(1.92)
$\beta_{HML^{FF}}$	0.50	0.52	0.43	0.56
	(19.46)	(13.24)	(9.46)	(11.93)
Adj. R^2	58.04	57.12	56.67	60.79
RMSE	4.23	4.07	4.47	4.12
α/RMSE	0.91	0.91	1.38	0.56
B. $\mathbf{HML_t^{FF}} = \boldsymbol{\alpha} + \boldsymbol{\beta_{HML^{INT}}} \cdot \mathbf{HML_t^{INT}} + \boldsymbol{\epsilon_t}$				
α (%)	-3.03	-1.88	-5.10	-3.59
	(-3.00)	(-1.30)	(-2.14)	(-2.18)
$\beta_{HML^{INT}}$	1.16	1.11	1.31	1.08
	(23.33)	(16.96)	(13.49)	(11.21)
Adj. R^2	58.04	57.12	56.67	60.79
RMSE	6.45	5.95	7.77	5.71
α/RMSE	-0.47	-0.32	-0.66	-0.63

Table 2.6: Single Factor Models – Intangible Value vs. Traditional Value

Description: In this table, we study the relative performance of the HML^{FF} and HML^{INT} factors. We report alphas and betas of a regression of each return on the other, for the full sample as well as for sub-periods around the Internet Bubble and the Great Recession. The data are monthly and the sample period is 1975 to 2018. We include T-statistics that adjust for heteroskedasticity in parentheses. All factors are annualized in percent per year.
Interpretation: The alpha of HML^{INT} over HML^{FF} is highly positive and significant. The positive alpha is also fairly stable over time and is significant in all subsamples.

	Full sample	1975-1994	1995-2006	2007-2018
	(1)	(2)	(3)	(4)
A. $\mathbf{HML}_t^{\mathbf{INT}} = \boldsymbol{\alpha} + \boldsymbol{\beta}_{\boldsymbol{HML}^{\mathbf{PTINT}}} \cdot \mathbf{HML}_t^{\mathbf{PTINT}} + \boldsymbol{\epsilon}_t$				
α (%)	2.42	1.78	4.60	1.72
	(4.41)	(2.75)	(4.18)	(1.42)
$\beta_{HML^{\text{PTINT}}}$	0.60	0.64	0.52	0.69
	(25.80)	(20.96)	(13.11)	(11.99)
Adj. R^2	70.12	77.65	70.69	59.36
RMSE	3.57	2.93	3.67	4.19
α/RMSE	0.68	0.61	1.25	0.41
B. $\mathbf{HML}_t^{\mathbf{PTINT}} = \boldsymbol{\alpha} + \boldsymbol{\beta}_{\boldsymbol{HML}^{\mathbf{INT}}} \cdot \mathbf{HML}_t^{\mathbf{INT}} + \boldsymbol{\epsilon}_t$				
α (%)	-1.24	-0.57	-3.67	-2.06
	(-1.66)	(-0.59)	(-2.18)	(-1.53)
$\beta_{HML^{\text{INT}}}$	1.16	1.21	1.36	0.87
	(26.17)	(23.79)	(18.16)	(11.98)
Adj. R^2	70.12	77.65	70.69	59.36
RMSE	4.94	4.04	5.91	4.70
α/RMSE	-0.25	-0.14	-0.62	-0.44

Table 2.7: Single Factor Models – Alternative Intangible Asset Calculation Methods

Description: In this table, we study the relative performance of our baseline HML^{INT} and $\text{HML}^{\text{PTINT}}$, the factor that accumulates 30% of (SG&A-R&D) plus 100% of R&D to construct intangible assets and sort firms across all industries (see Online Appendix for details). We report alphas and betas of a regression of each return on the other, for the full sample as well as for sub-periods around the Internet Bubble and the Great Recession. The data are monthly and the sample period is 1975 to 2018. We include T-statistics that adjust for heteroskedasticity in parentheses. All factors are annualized in percent per year.
Interpretation: Our HML^{INT} factor has a positive alpha over $\text{HML}^{\text{PTINT}}$ in the full sample and in all subsamples except the post-crisis era. We conclude that our intangible value factor outperforms the factor used in Park (Forthcoming).

	Full sample	1975-1994	1995-2006	2007-2018
	(1)	(2)	(3)	(4)
A. $\mathbf{HML_t^{IME}} = \boldsymbol{\alpha} + \boldsymbol{\beta_{HML^{FF}}} \cdot \mathbf{HML_t^{FF}} + \boldsymbol{\epsilon_t}$				
α (%)	4.95	4.65	6.99	3.41
	(7.03)	(4.71)	(5.21)	(2.42)
$\beta_{HML^{FF}}$	0.40	0.46	0.33	0.41
	(14.23)	(11.43)	(6.59)	(7.19)
Adj. R^2	41.59	45.17	40.60	36.35
RMSE	4.73	4.60	4.70	4.86
α/RMSE	1.05	1.01	1.49	0.70
B. $\mathbf{HML_t^{UINT}} = \boldsymbol{\alpha} + \boldsymbol{\beta_{HML^{FF}}} \cdot \mathbf{HML_t^{FF}} + \boldsymbol{\epsilon_t}$				
α (%)	4.71	3.46	7.69	6.19
	(2.85)	(1.35)	(2.36)	(2.25)
$\beta_{HML^{FF}}$	-0.07	-0.07	-0.28	0.27
	(-1.09)	(-0.59)	(-2.56)	(2.96)
Adj. R^2	0.25	-0.10	7.63	5.77
RMSE	10.91	11.36	10.97	9.41
α/RMSE	0.43	0.30	0.70	0.66

Table 2.8: Single Factor Models – Decompositions of Intangible Value

Description: In this table, we report alphas and betas of a regression of HML^{IME} and HML^{UINT} on HML^{FF}, for the full sample as well as for sub-periods around the Internet Bubble and the Great Recession. HML^{IME} is constructed using the intangible capital-to-market value ratio as the sorting variable. HML^{UINT} is a portfolio that is long firms that are sorted in the long leg when using $\text{B}^{\text{INT}}/\text{M}$ but not when using B/M, and similarly, short firms that are uniquely in the short leg of HML^{INT}. The data are monthly and the sample period is 1975 to 2018. We include T-statistics that adjust for heteroskedasticity in parentheses. All factors are annualized in percent per year.
Interpretation: The two portfolios that isolate the effect of intangibles display more independent variation relative to traditional value (lower R^2). Moreover, the full sample alphas for HML^{IME} and HML^{UINT} are higher than the alpha for HML^{INT} (Table 2.6, Panel A), supporting the return-enhancing effect of including intangibles in value anchors.

	Full sample	1975-1994	1995-2006	2007-2018
	(1)	(2)	(3)	(4)
$\mathbf{HML}_t^{\mathbf{INT}} = \boldsymbol{\alpha} + \boldsymbol{\beta}_{\boldsymbol{OMK}} \cdot \mathbf{OMK}_t + \boldsymbol{\epsilon}_t$				
α (%)	5.51	6.47	8.29	0.99
	(5.65)	(4.60)	(4.86)	(0.55)
β_{OMK}	0.04	-0.05	0.25	-0.27
	(0.79)	(-0.74)	(4.30)	(-2.93)
Adj. R^2	0.09	0.04	18.72	10.21
RMSE	6.52	6.21	6.12	6.23
α/RMSE	0.85	1.04	1.36	0.16

Table 2.9: Single Factor Models – Intangible Value and Organization Capital Factor

Description: In this table, we report alphas and betas of a regression of HML^{INT} on the OMK factor (Eisfeldt and Papanikolaou, 2013b), for the full sample as well as for sub-periods around the Internet Bubble and the Great Recession. The data are monthly and the sample period is 1975 to 2018. We include T-statistics that adjust for heteroskedasticity in parentheses. All factors are annualized in percent per year.

Interpretation: The R^2 in a regression of intangible value on the organization capital factor from Eisfeldt and Papanikolaou (2013b) is negligible, implying that the two factors capture different effects of intangibles both conceptually and empirically.

		Full sample	1975-1994	1995-2006	2007-2018
		(1)	(2)	(3)	(4)
	$\mathbb{E}[R]$	3.49	5.14	6.99	-2.77
		(2.33)	(2.53)	(2.05)	(-1.05)
HML$^{\mathbf{FF}}$	σ	9.95	9.08	11.80	9.11
	[0.05, 0.95]	[-48.36, 63.24]	[-45.72, 63.12]	[-55.92, 78.24]	[-44.04, 48.84]
	Sharpe	0.35	0.57	0.59	-0.30
	$\mathbb{E}[R]$	5.60	6.34	9.21	0.76
		(5.70)	(4.57)	(4.70)	(0.40)
HML$^{\mathbf{INT}}$	σ	6.52	6.21	6.78	6.57
	[0.05, 0.95]	[-27.54, 40.43]	[-23.63, 40.17]	[-25.95, 48.42]	[-36.38, 35.93]
	Sharpe	0.86	1.02	1.36	0.12
	$\mathbb{E}[R]$	6.35	7.02	9.30	2.28
		(6.81)	(5.06)	(5.28)	(1.30)
HML$^{\mathbf{IME}}$	σ	6.18	6.21	6.10	6.09
	[0.05, 0.95]	[-26.48, 40.80]	[-25.11, 40.98]	[-20.31, 45.42]	[-35.03, 36.87]
	Sharpe	1.03	1.13	1.53	0.37
	$\mathbb{E}[R]$	2.11	1.20	2.22	3.53
		(2.15)	(0.90)	(0.96)	(2.14)
HML$^{\mathbf{INT}}$	σ	6.53	5.97	8.03	5.71
- HML$^{\mathbf{FF}}$	[0.05, 0.95]	[-36.59, 36.54]	[-32.60, 34.39]	[-44.07, 45.00]	[-26.31, 30.74]
	Information	0.32	0.20	0.28	0.62
	Appraisal	0.91	0.91	1.38	0.56
	$\mathbb{E}[R]$	2.86	1.88	2.31	5.05
		(2.50)	(1.25)	(0.87)	(2.41)
HML$^{\mathbf{IME}}$	σ	7.60	6.71	9.18	7.27
- HML$^{\mathbf{FF}}$	[0.05, 0.95]	[-40.67, 43.14]	[-39.18, 39.02]	[-50.84, 53.19]	[-37.68, 41.31]
	Information	0.38	0.28	0.25	0.70
	Appraisal	1.05	1.01	1.49	0.70

Table 2.10: Performance Statistics – Intangible Value vs. Traditional Value

Description: This table summarizes the risk and return associated with intangible and traditional value. HML$^{\text{INT}}$−HML$^{\text{FF}}$ refers to the portfolio that is long HML$^{\text{INT}}$ and short HML$^{\text{FF}}$, and HML$^{\text{IME}}$−HML$^{\text{FF}}$ refers to the portfolio that is long HML$^{\text{IME}}$ and short HML$^{\text{FF}}$. The numbers in parentheses are T-statistics for the test that the average return, $\mathbb{E}[R]$, is different from zero. The information ratio is $\mathbb{E}[R_p - R_b]/\sigma(R_p - R_b)$, or the Sharpe Ratio of the long-short portfolio. The appraisal ratio is α/RMSE of a regression of intangible value returns (HML$^{\text{INT}}$ or HML$^{\text{IME}}$) on traditional value returns. The underlying data are monthly and the full sample period is 1975 to 2018. All factors are annualized in percent per year.
Interpretation: HML$^{\text{INT}}$ and HML$^{\text{IME}}$ consistently exhibits higher returns and lower volatility than HML$^{\text{FF}}$, resulting in a higher Sharpe ratio for all periods. A portfolio that is long intangible value (either variant) and short traditional value has positive returns in the full sample and in the post-crisis era when all value factors underperformed.

	HML^{FF}	HML^{INT}	HML^{IME}	HML^{UINT}	HML^{FF}	HML^{INT}	HML^{IME}	HML^{UINT}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α (%)	-1.43	2.92	3.87	2.81	-1.67	2.15	3.15	1.27
	(-1.51)	(4.93)	(5.67)	(1.82)	(-1.84)	(3.64)	(4.69)	(0.82)
β_{MktRF}	-0.09	0.04	0.04	0.06	-0.05	0.06	0.07	0.10
	(-5.26)	(2.96)	(3.17)	(1.55)	(-2.36)	(4.92)	(5.00)	(2.69)
β_{SMB}	-0.27	0.19	0.19	0.51	-0.23	0.21	0.20	0.58
	(-9.62)	(10.98)	(9.68)	(11.72)	(-7.70)	(10.81)	(9.47)	(11.77)
$\beta_{HML^{INT}}$	1.20				0.97			
	(26.12)				(15.63)			
$\beta_{HML^{FF}}$		0.55	0.46	0.05		0.46	0.35	-0.09
		(26.42)	(17.71)	(0.95)		(17.56)	(11.13)	(-1.35)
β_{UMD}	-0.05	-0.01	0.00	-0.05	-0.06	-0.02	-0.01	-0.07
	(-1.87)	(-0.33)	(0.12)	(-1.21)	(-2.83)	(-1.26)	(-0.88)	(-2.02)
β_{RMW}					0.00	0.09	0.06	0.23
					(0.04)	(3.45)	(1.82)	(3.33)
β_{CMA}					0.36	0.20	0.24	0.25
					(6.20)	(5.39)	(5.13)	(2.65)
Adj. R^2	70.57	68.46	53.98	25.47	73.50	70.90	57.17	28.12
RMSE	5.40	3.66	4.19	9.43	5.12	3.52	4.05	9.27

Table 2.11: Alphas – Intangible Value vs. Traditional Value

Description: In this table, we report portfolio alphas and betas of a regression of different variants of HML portfolio returns on traditional factor models. Columns (1) through (4) use the Fama and French (1992, 1993) three factor model, and columns (5) through (8) use the Fama and French (2015) five factor model. Both specifications are augmented with the momentum factor. Columns (1) and (5) are benchmarks that set HML^{IME} as the dependent variable and replace the intangibles-adjusted HML factor in the aforementioned models. We include T-statistics that adjust for heteroskedasticity in parentheses. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve).
Interpretation: The intangible value factors (HML^{INT} and HML^{IME}) have positive and significant alphas in the three- and five-factor models plus momentum, with the exception of HML^{UINT}, for which the positive alphas are insignificant. Alphas from the traditional factor models are insignificant.

	B^{INT}/M			B/M		
	Low 30 (1)	Mid 40 (2)	High 30 (3)	Low 30 (4)	Mid 40 (5)	High 30 (6)
B/M Int	0.58	1.56	3.93	0.75	1.66	3.42
B/M FF	0.34	0.70	1.15	0.31	0.74	1.38
Market capitalization (log, real)	5.45	4.88	3.34	5.03	4.76	3.58
Intangible capital to book assets (%)	37.44	56.78	97.92	74.97	53.27	48.29
Intangible capital to sales (%)	43.15	60.13	87.10	72.64	59.49	62.18
Productivity - sales to book assets (%)	81.62	90.32	103.31	96.13	92.29	80.95
Productivity - Solow residual (%)	1.30	1.77	0.07	6.10	1.33	-5.51
Sales to Stockholder's equity (%)	175.72	188.92	227.94	205.61	194.92	186.85
Price to Earnings (Diluted, excluding extraordinary items) (%)	14.00	11.96	7.33	13.40	11.95	8.34
Price to sales (%)	1.91	1.11	0.62	1.86	0.97	0.59
Debt to book assets (%)	15.44	12.60	15.25	10.43	14.15	16.92
Debt to EBITDA (%)	77.00	141.84	172.95	45.98	159.27	231.85
Dividend yield	2.14	2.86	3.01	1.92	2.94	3.30
Investment to physical capital (%)	14.69	11.19	8.72	14.85	10.66	8.19
Gross profit to total assets (%)	26.69	27.77	30.53	37.88	27.51	19.19
Exposure to Technology Shocks	-0.015	-0.096	-0.182	-0.027	-0.093	-0.146

Table 2.12: Summary Statistics of Firm Characteristics

Description: This table summarizes the characteristics of firms sorted into the long ("High 30") and short ("Low 30") legs of the HML^{FF} and HML^{INT} factors. B/M is the traditional book-to-market ratio, and B^{INT}/M denotes the intangible-adjusted book-to-market ratio used to construct HML^{INT}. We report the time-series average of the median firm characteristic within each percentile bucket. The sample period is January 1975 to December 2018.

Interpretation: On average, firms in the long leg of intangible value have better productivity, performance, leverage ratios, and alternative valuation ratios compared to firms in the short leg. The relationship is flipped for average characteristics of firms in the long and short legs of traditional value. The table supports the idea that intangible value tends to better avoid value traps and avoids shorting firms with valuable, but unmeasured, intangible assets.

j	High^{INT}	Low^{INT}	High^{FF}	Low^{FF}
$\mathbf{P}$	$\begin{bmatrix} 57.87 & 42.13 \\ 8.00 & 92.00 \end{bmatrix}$	$\begin{bmatrix} 54.23 & 45.77 \\ 3.57 & 96.43 \end{bmatrix}$	$\begin{bmatrix} 51.49 & 48.51 \\ 4.19 & 95.81 \end{bmatrix}$	$\begin{bmatrix} 57.51 & 42.49 \\ 6.90 & 93.10 \end{bmatrix}$
$\mathbf{w}$	$(15.97, 84.03)$	$(7.23, 92.77)$	$(7.95, 92.05)$	$(13.97, 87.03)$

Table 2.13: Persistence of Positions

Description: This table represents transition matrices $\mathbf{P}$ for being sorted uniquely into a particular leg of the HML^{INT} and HML^{FF} portfolios. For instance, the state $j = \text{High}^{\text{INT}}$ refers to a given firm being sorted in the top 30th percentile in terms of $\text{B}^{\text{INT}}/\text{M}$ and in the bottom 70th percentile in terms of B/M. In this case, the alternative state can be either i) being sorted in the top 30th percentiles of both $\text{B}^{\text{INT}}/\text{M}$ and B/M, or ii) being sorted in the bottom 70th percentile of $\text{B}^{\text{INT}}/\text{M}$, regardless of the B/M sort. Below each panel, we report the stationary distribution, $\mathbf{w} = (\pi_j, 1 - \pi_j)$, of each Markov Chain, where π_j denotes the long run proportion of time that each chain spends in state j. All numbers are expressed in percentages.
Interpretation: Unique assignment into each leg of the intangible or traditional value factor (e.g. long in intangible sort but not so in traditional sort) is persistent. This indicates that turnover of portfolio constituents is unlikely to be the driver of differences between HML^{INT} and HML^{FF}. Moreover, intangible value turnover is similar to that for traditional value; both are slow-moving factors.

2.B Data Appendix

Constructing $\mathrm{HML}^{\mathrm{INT}}$ involves a three-step process: First, we calculate the firm-level stock of intangibles using the perpetual inventory method. Next, we add intangibles to book value of equity and subtract goodwill. Lastly, we sort firms within industries based on their intangibles-augmented book-to-market ratio and form hedged long-short portfolios. In this section, we describe this process in further detail. The relevant code and programs are also posted on the authors' websites.

2.B.1 Measuring Intangible Capital: EKP Method

We compute a measure of book equity including intangibles using the following formula:

$$\mathrm{B}_{it}^{\mathrm{INT}} = \mathrm{B}_{it} - \mathrm{GDWL}_{it} + \mathrm{INT}_{it}, \tag{2.5}$$

where B_{it} is book equity, GDWL_{it} is goodwill (Compustat item *gdwl*), and INT_{it} is intangible assets for firm i at time t.[17]

To compute $\mathrm{B}_{it}^{\mathrm{INT}}$, we first calculate the stock of intangible assets at the firm-level using methodology based on Eisfeldt and Papanikolaou (2013b), and Eisfeldt and Papanikolaou (2013a), Eisfeldt and Papanikolaou (2014). Intangible assets created internally are expensed and typically do not appear explicitly on the balance sheet. This means that the replacement cost of internally generated intangible assets must be calculated based on past investments in intangibles. As this investment is also not measured and reported under standard accounting practices, we must find a proxy and accumulate this identity over time. Our preferred method follows the original method in Eisfeldt and Papanikolaou (2013b), which we denote in the context of intangible value by "EKP method". Using this method, we construct $\mathrm{B}_{it}^{\mathrm{INT}}$ using past

[17] Following Fama and French (1992, 1993), we calculate book equity using Compustat data: $be = (seq \text{ or } ceq + pstk \text{ or } at - lt) + (txditc \text{ or } txdb + itcb) + (pstkrv \text{ or } pstkl \text{ or } pstk)$

investments in selling, general, and administrative (SG&A) expenses (item *xsga*). Specifically, the perpetual inventory method allows for the stock of intangibles to grow with the law of motion:

$$\text{INT}_{it} = (1-\delta)\text{INT}_{it-1} + \text{SG\&A}_{it}. \tag{2.6}$$

where $\delta_{SG\&A}$ is the depreciation rate for SG&A expenses and $\text{SG\&A}_{it}$ is real SG&A expenditure, calculated by deflating *xsga* by the consumer price index. Moreover, we set $\text{INT}_{i0} = \text{SG\&A}_{i1}/(g+\delta)$ and use $g = 0.1$ to compute the initial stock of organization capital prior to the first observation in Compustat. Prior works including Eisfeldt and Papanikolaou (2013a) provide detailed justification for this procedure. For our analysis, we set $\delta = 0.2$, and in unreported results, we verify that using different values of reasonable depreciation rates do not meaningfully change our conclusions. Lastly, we apply this algorithm to all firms in Compustat from 1950 and begin our sample in 1975.

Intangible assets acquired through a purchase — for instance, by acquiring another firm — are capitalized on the balance sheet as either "Goodwill (item *gdwl*)" or "Other Intangible Assets (item *intano*)," the sum of which is readily available as item *intan*. *intan* is already incorporated into book assets (item *at*), so we do not add this variable to our measure of total assets accounting for intangibles. The goodwill component of *intan* arises when merger values exceed book values by more than the value of identifiable intangible assets, and reflects market values in excess of book values including identifiable intangibles at the time of the merger. We thus subtract goodwill from book equity.

2.B.2 Comparison to Alternative Intangible Capital Method: PT Method

In a robustness exercise ("PT method"), we follow Peters and Taylor (2017) that break down a firm's intangible capital (INT_{it}) into the sum of two components — *knowledge capital* (e.g. R&D spending) and *organization capital* (e.g. human capital, brand capital, and customer relationships). Here, we use the R&D (item xrd) and SG&A (item $xsga$) variables from Compustat to calculate INT^{know} and INT^{org}, respectively. Specifically, we estimate the following for INT^{know}

$$\text{INT}_{it}^{know} = (1 - \delta_{R\&D})\text{INT}_{it-1}^{know} + \text{R\&D}_{it}, \tag{2.7}$$

where INT_{it}^{know} is the stock of knowledge capital, $\delta_{R\&D}$ is an industry-specific depreciation rate for knowledge capital, and $\text{R\&D}_{it}$ is the real expenditures on R&D, which is measured by deflating Compustat item *xrd*. Data on industry-specific depreciation rates are obtained from Li and Hall (2020) and range from 10% to 40%.[18] We initialize $\text{INT}_{i0}^{know} = \text{R\&D}_{i1}/(g + \delta_{R\&D})$ where $g = 0.1$.

The book stock of organization capital, INT^{org}, can be similarly estimated by applying the law of motion

$$\text{INT}_{it}^{org} = (1 - \delta_{SG\&A})\text{INT}_{it-1}^{org} + \theta\text{SG\&A}_{it}, \tag{2.8}$$

where $\text{SG\&A}_{it}$ is real SG&A expenditure calculated by subtracting xrd from $xsga$ and deflating the resulting stock by the consumer price index. We subtract xrd from $xsga$ because xrd is included in $xsga$ under standard accounting practices. $\delta_{SG\&A}$ is the depreciation rate specific to SG&A expenses, which we assume is 0.2. θ is the investment rate for organization capital, which we set $\theta = 0.3$ following Peters and Taylor (2017). We initialize $\text{INT}_{i0}^{org} = \theta\text{SG\&A}_{i1}/(g + \delta_{SG\&A})$ where $g = 0.1$.

[18] We apply $\delta = 0.15$ for the majority of SIC codes that are not assigned a specific depreciation rate.

We verify that using different values of reasonable depreciation and investment rates do not meaningfully change our results. Finally, the PT measure of total intangible capital is calculated as

$$\text{PTINT}_{it} = \text{INT}_{it}^{know} + \text{INT}_{it}^{org}. \tag{2.9}$$

2.B.3 Intangible Value Factor

The key empirical goal of estimating intangible capital is to construct a modified book-to-market equity ratio, which is in turn used to form the Fama and French (1992, 1993) value factor. Book assets serve as a balance sheet benchmark for each firm's intrinsic value, and the ratio between this anchor and the market equity value measures the extent of over- or under-valuation. For our intangibles-adjusted measure of value, we divide $\text{B}_{it}^{\text{INT}}$ computed in Section 2.B.1 by the market value of equity, which is computed as *shrout* $\times$ *prc* using data from Center for Research in Security Prices (CRSP).

The intangible value factor is constructed using six annually rebalanced and value-weighted portfolios formed on size and $\text{B}^{\text{INT}}/\text{M}$. The six portfolios span the combination of two size (Small and Big with cutoff at median market capitalization) and three book-to-market (Value, Neutral, and Growth with book-to-market ratios in the top 30th percentile, between the 30th and 70th percentiles, and the bottom 30th percentile, respectively) portfolios. The *value factor*, commonly abbreviated as HML (High Minus Low), is the average return on the two value portfolios minus the average return on the two growth portfolios. Notably, unlike other works in the literature, we first compute a within-industry measure of HML

$$\text{HML}_{It} = \frac{1}{2}\left(\text{Small Value}_{It} + \text{Big Value}_{It}\right) - \frac{1}{2}\left(\text{Small Growth}_{It} + \text{Big Growth}_{It}\right), \tag{2.10}$$

where stock returns are measured monthly and I refers to each of the 12 industries classified by Fama and French. Then we compute $\mathrm{HML}^{\mathrm{INT}}$ as

$$\mathrm{HML}_t^{\mathrm{INT}} = \sum_{I=1}^{12} w_{It} \times \mathrm{HML}_{It}, \tag{2.11}$$

where w_{It} is the weight of each industry's total market capitalization. While common in the literature, we do not drop industries such as financials or regulated utilities for our intangible value factor in order to ensure that our method replicates the original Fama and French method as closely as possible. The PT method follows this procedure, the only distinction being the use of $\mathrm{B}^{\mathrm{PTINT}}$ in the numerator of the B/M ratio.

2.B.4 Other Measures of Intangible Value

For our main analyses, we additionally study various alternative measures of intangible value in order to analyze the unique pricing ability of $\mathrm{HML}^{\mathrm{INT}}$.

First, $\mathrm{HML}^{\mathrm{IME}}$ is a value factor that sorts firms into high and low buckets based on INT/ME instead of $\mathrm{B}^{\mathrm{INT}}$/M. This factor isolates the portion of value that is purely attributable to intangible assets. Specifically, we define Value as high-INT/ME and Growth as low-INT/ME and construct six annually rebalanced portfolios for each industry I following the EKP method

$$\mathrm{HML}_{It}^{\mathrm{IME}} = \frac{1}{2}\left(\text{Small Value}_{It} + \text{Big Value}_{It}\right) - \frac{1}{2}\left(\text{Small Growth}_{It} + \text{Big Growth}_{It}\right). \tag{2.12}$$

The IME factor construction process is also consistent with the EKP method

$$\mathrm{HML}_t^{\mathrm{IME}} = \sum_{I=1}^{12} w_{It} \times \mathrm{HML}_{It}^{\mathrm{IME}}, \tag{2.13}$$

We also introduce HML^{UINT}, which sorts firms on $\text{B}^{\text{INT}}/\text{M}$ but only goes long firms that are *uniquely* in the long leg of HML^{INT} (i.e. not sorted in the long leg of HML^{FF}), and goes short firms that are *uniquely* in the short leg of HML^{INT} (i.e. not sorted in the short leg of HML^{FF}). To construct HML^{UINT}, we identify "unique long" firms as those above the 70th percentile in $\text{B}^{\text{INT}}/\text{M}$ but below the 70th percentile in the distribution of B/M across all industries. An analagous approach is used to identify the "unique short" firms. After identifying this subset of firms, we value-weight the returns of each stock in each leg and construct the long-short portfolio:

$$\text{HML}_t^{\text{UINT}} = \sum_{i=1}^{n} w_{it} \times \text{Unique Long}_{it} - \sum_{j=1}^{m} w_{jt} \times \text{Unique Short}_{jt}. \tag{2.14}$$

Note that HML^{UINT} is not sorted within industries and industry-weighted in the second step because of the lower number of firms included in each leg. For this process, we adhere to the simple sorting and portfolio formation methodology that mimics Fama and French (1992, 1993).

INT-FF is a factor that is simply HML^{INT} minus HML^{FF}. Similarly, IME-FF is HML^{IME} minus HML^{FF}. For these two factors, note that there may be firms sorted into the same long-short legs but with different portfolio weights. We assume an investor can passively buy HML^{INT} (or HML^{IME}) and sell HML^{FF} in exactly offsetting amounts. Moreover, we construct $\text{HML}^{\text{INDFF}}$, which is the Fama and French HML factor that follows our within-industry sorting and weighting methodology.

Lastly, we also create a version of HML^{INT} that drops financials (SIC codes 6,000-6,999), regulated utilities (4,900-4,999), and firms categorized as public service, international affairs, or non-operating establishments (9,000+).

2.C Further Analysis and Robustness Checks

In this section, we study the relative performance of the long and short legs of HML^{INT} and HML^{FF}, and report our main results using various robustness measures of value.

2.C.1 Further Long and Short Leg Analysis

In this section, we study the relative performance of the long and short legs of HML^{INT} and HML^{FF}. For H^{INT} and L^{INT}, we compute the returns of the long and short leg for each industry, and weight those industry leg returns by industry market cap. H^{FF} and L^{FF} are obtained from Ken French's website. The top panel of Figure 2.5 shows that on net, the cumulative returns of the long leg of intangible value is higher than the returns of traditional value's long leg. Similarly, the short leg of HML^{INT} consistently underperforms the short leg of HML^{FF}, meaning that the short side of the intangible value strategy is also more profitable (Figure 2.5, bottom panel). These results together show that the outperformance of intangible value is coming from both the long and short legs, and are not driven by a single leg. However, the long leg's outperformance is more pronounced starting in the 2010s while the short leg's outperformance begins earlier in the 1990s.

2.C.2 12 Industry Sorts for Traditional Value

In this section, we test whether our main asset pricing and performance results are driven by the within-industry sorting method. As noted in Section 2.1, we employ two crucial innovations to calculate our value factor – incorporating intangible capital to book value and sorting firms within industries. In this exercise, we replicate the original Fama and French HML factor (full-sample correlation of 98.0%) and create a within-industry sorted version, $\text{HML}^{\text{INDFF}}$. We compare $\text{HML}^{\text{INDFF}}$ to HML^{INT} and reproduce the main results below.

First, we examine the relationship between HML^{INT} and $\text{HML}^{\text{INDFF}}$. Figure 2.6 shows that the full-period correlation between returns of the two series is 0.89, which is markedly higher than the 0.76 correlation we reported in Figure 2.1 using HML^{FF}. In Figure 2.7, we see that the correlation between an unconditionally sorted HML^{INT} and unconditionally sorted HML^{FF} is 0.79. Taken together, both incorporating intangibles *and* sorting firms within industries help provide the variation in our baseline HML^{INT} series.

We reproduce our main regression results and compare the industry-sorted HML^{INT} to industry-sorted HML^{FF}. First, Table 2.D1 shows that industry-adjustment improves the asset pricing performance of $\text{HML}^{\text{INDFF}}$ as seen in the reduction of root mean squared errors in Columns (1) and (3). Moreover, the mean absolute pricing error of the three-factor model plus momentum in Figure 2.8 is noticeably reduced when using $\text{HML}^{\text{INDFF}}$. This is to be expected given the higher correlation between the $\text{HML}^{\text{INDFF}}$ and HML^{INT}. Despite this, the results are consistent with our observation that HML^{INT} prices assets as well as or better than HML^{FF} or $\text{HML}^{\text{INDFF}}$.

Table 2.D2 shows single factor models that test the outperformance of HML^{INT} over $\text{HML}^{\text{INDFF}}$. While the magnitude is slightly lower, the alpha of HML^{INT} over $\text{HML}^{\text{INDFF}}$ is positive and highly significant (2.16% vs. 3.86% for the baseline using HML^{FF}), consistent with findings in Table 2.6. Summary statistics on factor returns (Table 2.D3) also confirm that returns of $\text{HML}^{\text{INDFF}}$ are marginally improved when employing the within-industry sorting and weighting methodology (4.06% vs 3.49% for the full sample).

Table 2.D4 displays alphas of the traditional and intangible value factors in the three- and five-factor models plus momentum. We include results for the baseline intangible value factor, and for the two factors that isolate the effect of intangible capital. The alphas for industry-sorted traditional value (Columns (1) and (5)) are negative as in Table 2.11. For both models, the alpha for HML^{INT} is positive and sig-

nificant. The alphas for HML^{IME} are also positive and significant under both models. The intangible value factors all have positive and significant alphas in the three- and five-factor models with momentum, with the exception of HML^{UINT}, for which the positive alpha in the three-factor model is not significant.

2.C.3 Industry Filters

In this section, we report our main results after dropping financial firms (SIC codes 6,000-6,999), regulated utilities (4,900-4,999), and firms categorized as public service, international affairs, or non-operating establishments (9,000+), as is common in the literature. As our factor construction methodology accounts for industry differences, these filters likely only affect the relative weighting of the remaining industries' HML factors.

Table 2.D5 reproduces the baseline asset pricing test results dropping financials, utilities, and public service firms from the sample. While in general the alphas in models using intangible value are similar to or marginally higher than reported in Table 2.3, we find that dropping these industries do not materially change the pricing results. In particular, for the three-factor model with momentum, replacing the traditional value factor with the intangible value factor reduces both the alpha and root mean squared error. For the five-factor model with momentum, the alpha and root mean squared error under the two versions of value are largely analogous to results in Table 2.3.

Table 2.D6 shows single factor models that test the outperformance of intangible value relative to traditional value. Consistent with the main results in Table 2.6, the alpha of HML^{INT} over HML^{FF} is highly significant for the full sample and earlier sub-periods even after applying the industry filter. In fact, the magnitude of the alphas are notably higher when dropping these industries (e.g. 4.66% vs 3.86% for the full sample). These results are further corroborated by the improved performance

statistics of HML^{INT}, HML^{IME}, HML^{INT}-HML^{FF}, and HML^{IME}-HML^{FF} in Table 2.D7. Figure 2.11 visually shows the marked outperformance of HML^{INT} (solid blue line in top and bottom panels) when applying the industry filters. While the R^2 drop slightly, the portfolio alphas and betas reported in Table 2.D8 are also mostly unchanged.

2.D Appendix Figures and Tables

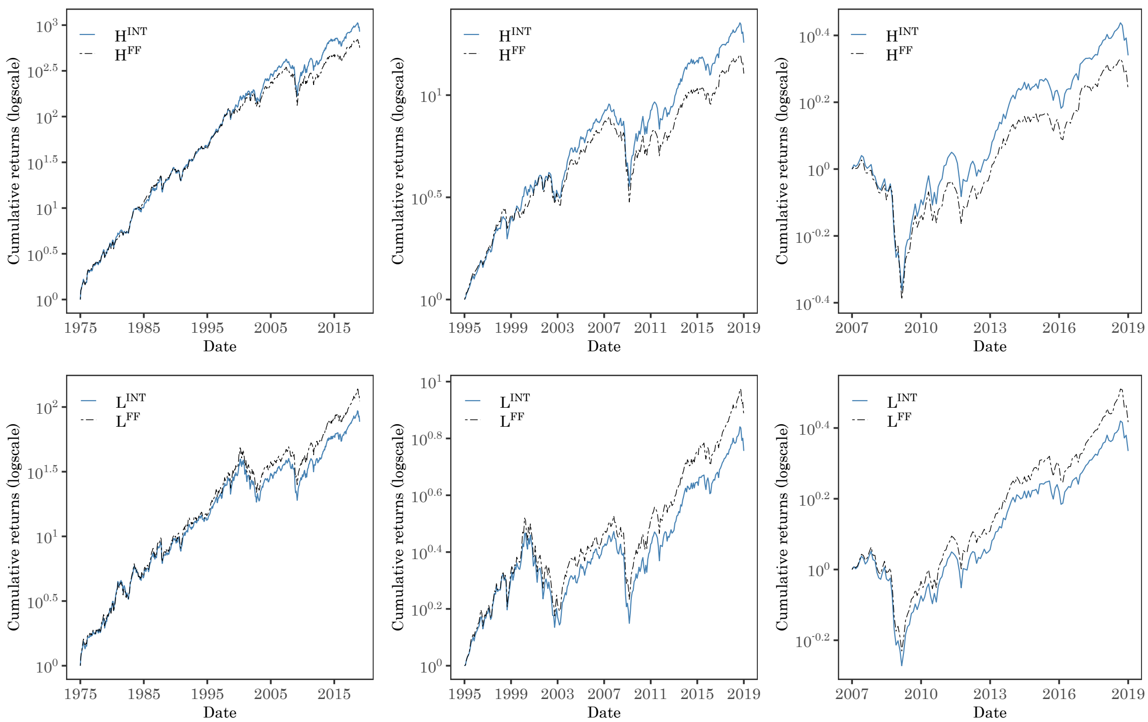


Figure 2.5: Performance of Long and Short Legs

Description: The top panel plots cumulative returns of the long leg of HML^{INT} (solid blue line) and the long leg of HML^{FF} (dashed black line). In the bottom panel, we plot the cumulative returns of the short leg of HML^{INT} (solid blue line) and the short leg of HML^{FF} (dashed black line). Each panel plots on a dollar invested in each leg from the beginning of 1975, 1995, and 2007.
Interpretation: Intangible value's outperformance arises from both the long and short legs.

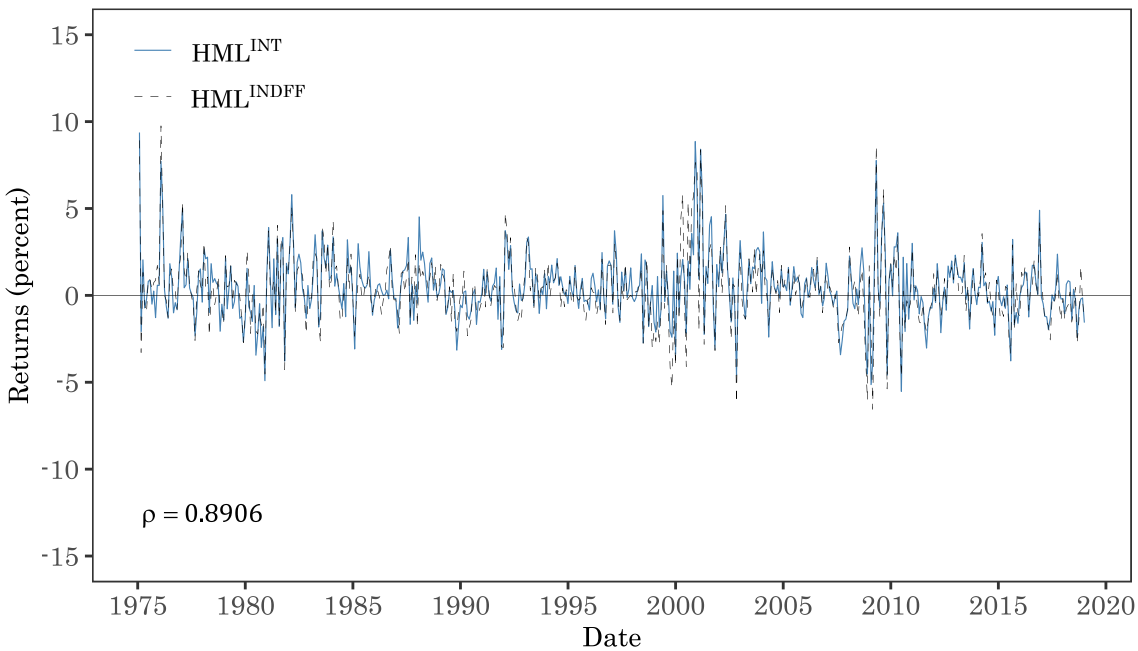


Figure 2.6: Traditional Value Sorted Within Industries

Description: This figure plots the monthly returns for HML^{INDFF} and HML^{INT} from 1975 to 2018. Firms are sorted within industries for both factors. The HML^{FF} portfolio mimics the risk factor in returns related to book-to-market equity, and is calculated as the difference between the returns on high-B/M portfolios and the returns on low-B/M portfolios.
Interpretation: As expected, sorting traditional value within industries increases the correlation between intangible value and traditional value.

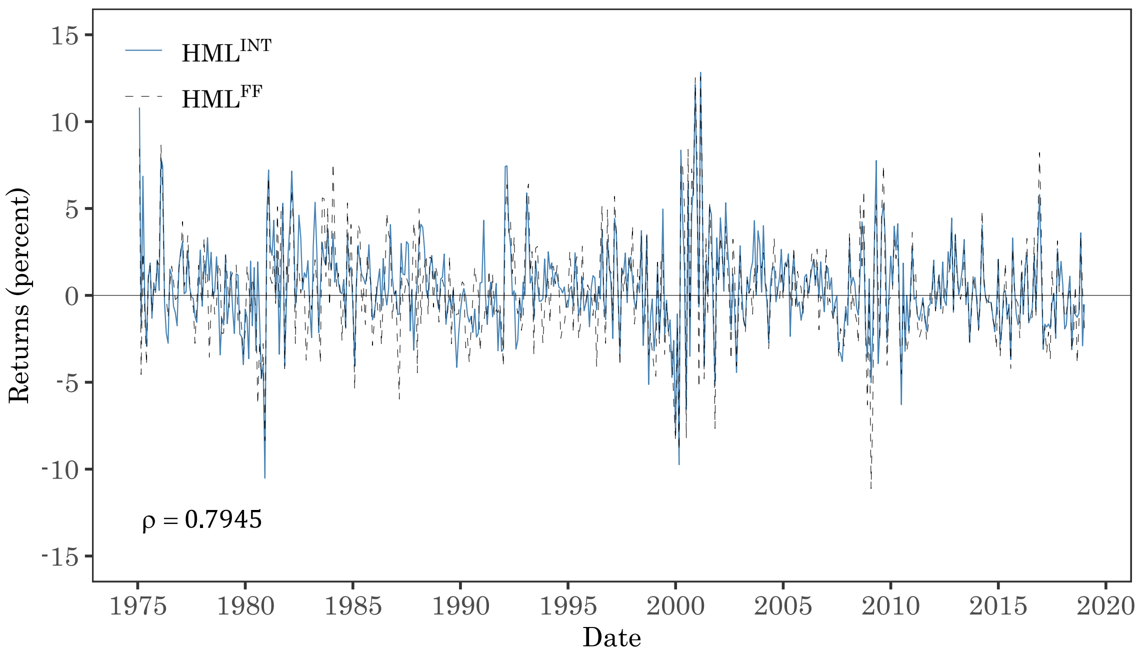


Figure 2.7: Intangible Value Sorted Across Industries

Description: This figure plots the monthly returns for HML^{FF} and HML^{INT} from 1975 to 2018. Firms are sorted unconditionally across industries for both factors. The HML^{FF} portfolio mimics the risk factor in returns related to book-to-market equity, and is calculated as the difference between the returns on high-B/M portfolios and the returns on low-B/M portfolios.
Interpretation: As expected, sorting intangible value across industries following the Fama and French methodology increases the correlation between intangible value and traditional value.

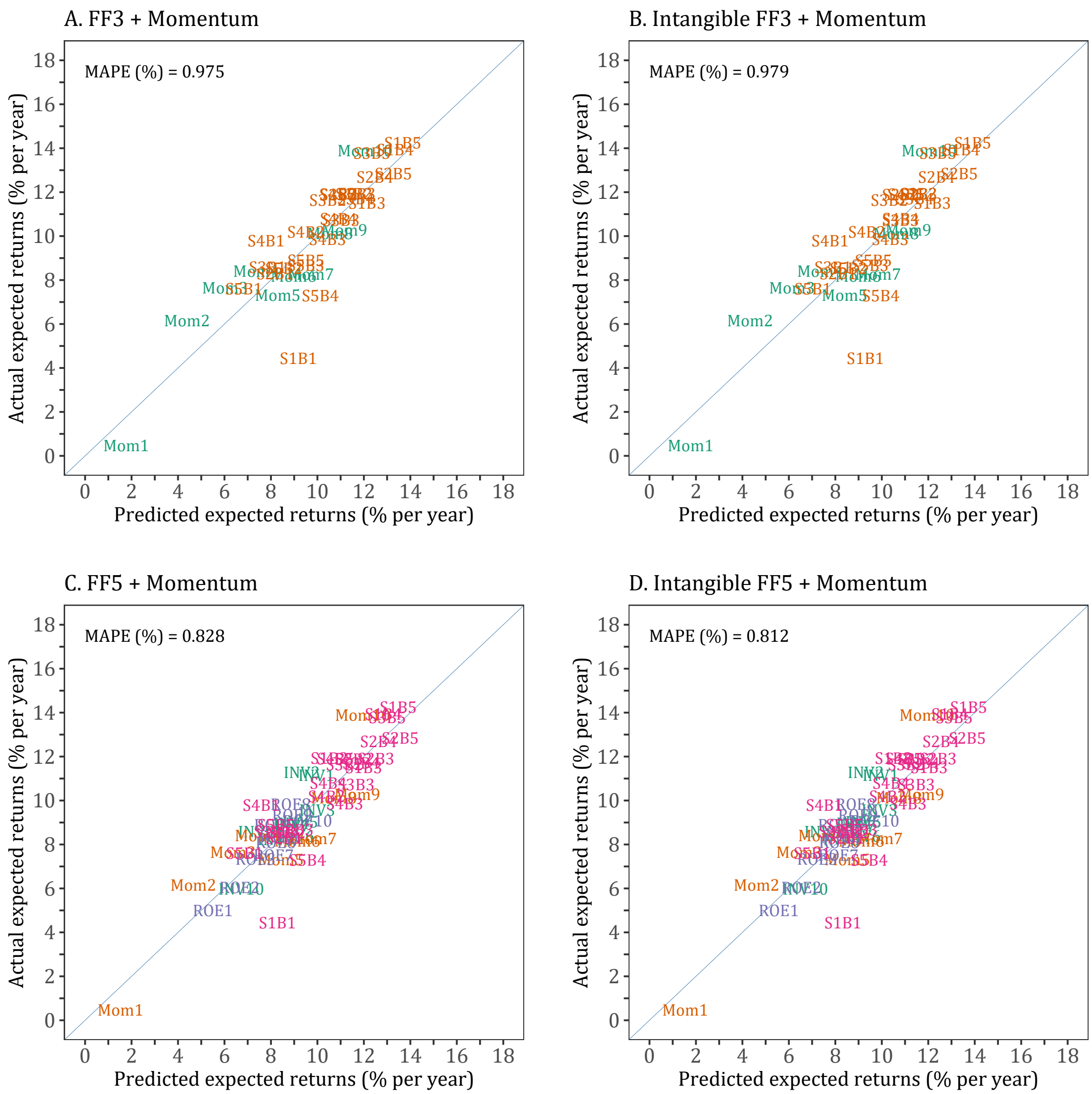


Figure 2.8: Cross-sectional Asset Pricing Tests – Industry-sorted Traditional Value

Description: This figure shows the cross-sectional asset pricing tests from the Fama and French (1992, 1993, 2015) three-factor and five-factor models augmented by the momentum factor. The top row plots realized mean excess returns of 25 size and book-to-market-sorted portfolios and 10 momentum portfolios against the mean excess returns predicted by the FF3 + momentum model, where Panel B replaces HML^{INDFF} with HML^{INT}. Firms are sorted within industries for both factors. The bottom row plots realized mean excess returns of 25 size and book-to-market-sorted portfolios, 10 momentum portfolios, 10 portfolios sorted on operating profitability, and 10 portfolios sorted on investment, against the mean excess returns predicted by the FF5 + momentum model. The sample is monthly from 1975 to 2018. Returns are reported in percent per year.

Interpretation: Sorting firms within industries improves the asset pricing performance of traditional value.

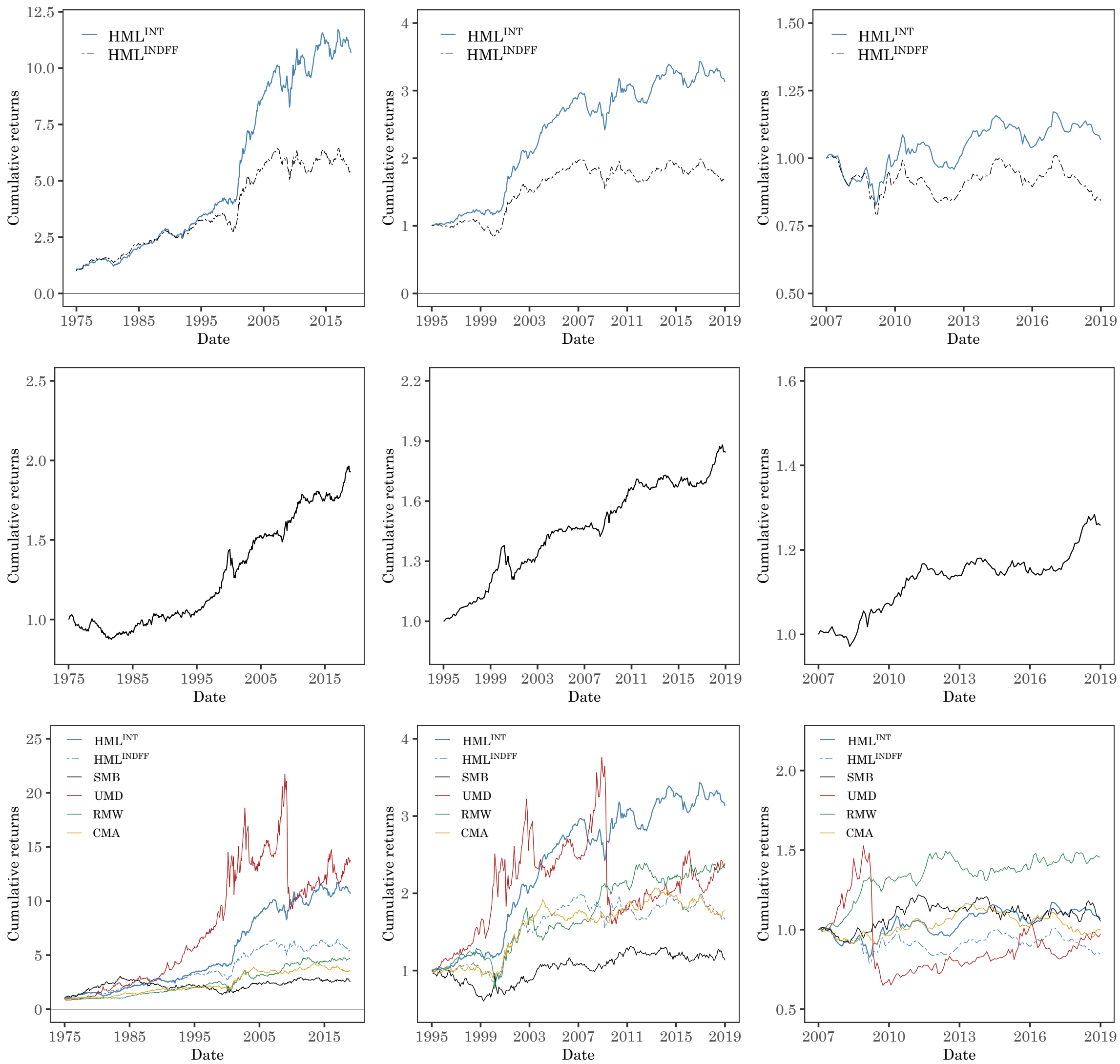


Figure 2.9: Performance of Industry-sorted Traditional Value

Description: The top panel plots the cumulative returns of one dollar invested in the HML$^{\text{INDFF}}$ and HML$^{\text{INT}}$ portfolios from the beginning of 1975, 1995, and 2007. The middle panel plots the cumulative returns of one dollar invested in the portfolio that is long the HML$^{\text{INT}}$ portfolio and short the HML$^{\text{INDFF}}$ portfolio. The bottom panel plots the cumulative returns of one dollar invested in HML$^{\text{INT}}$, the Fama and French five factors, and momentum.
Interpretation: Intangible value outperforms traditional value in both the full sample period and recent sub-samples. A long-short portfolio of intangible and traditional value also has positive returns. Lastly, intangible value exhibits similar performance as the top-performing momentum factor without suffering from the drawdown in the post-crisis era.

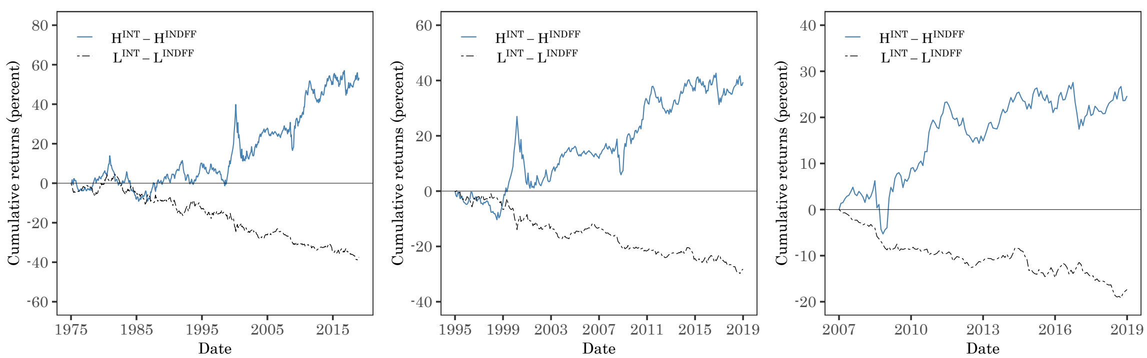


Figure 2.10: Decomposing Outperformance with Industry-sorted Traditional Value

Description: This figure plots the cumulative returns of a portfolio that is long the long leg of HML^{INT} and short the long leg of $\text{HML}^{\text{INDFF}}$ (solid blue line), as well as the returns of a portfolio that is long the short leg of HML^{INT} and short the short leg of $\text{HML}^{\text{INDFF}}$ (dashed black line). Each panel plots percent returns from the beginning of 1975, 1995, and 2007.
Interpretation: Each leg of intangible value outperforms traditional value that is industry-sorted.

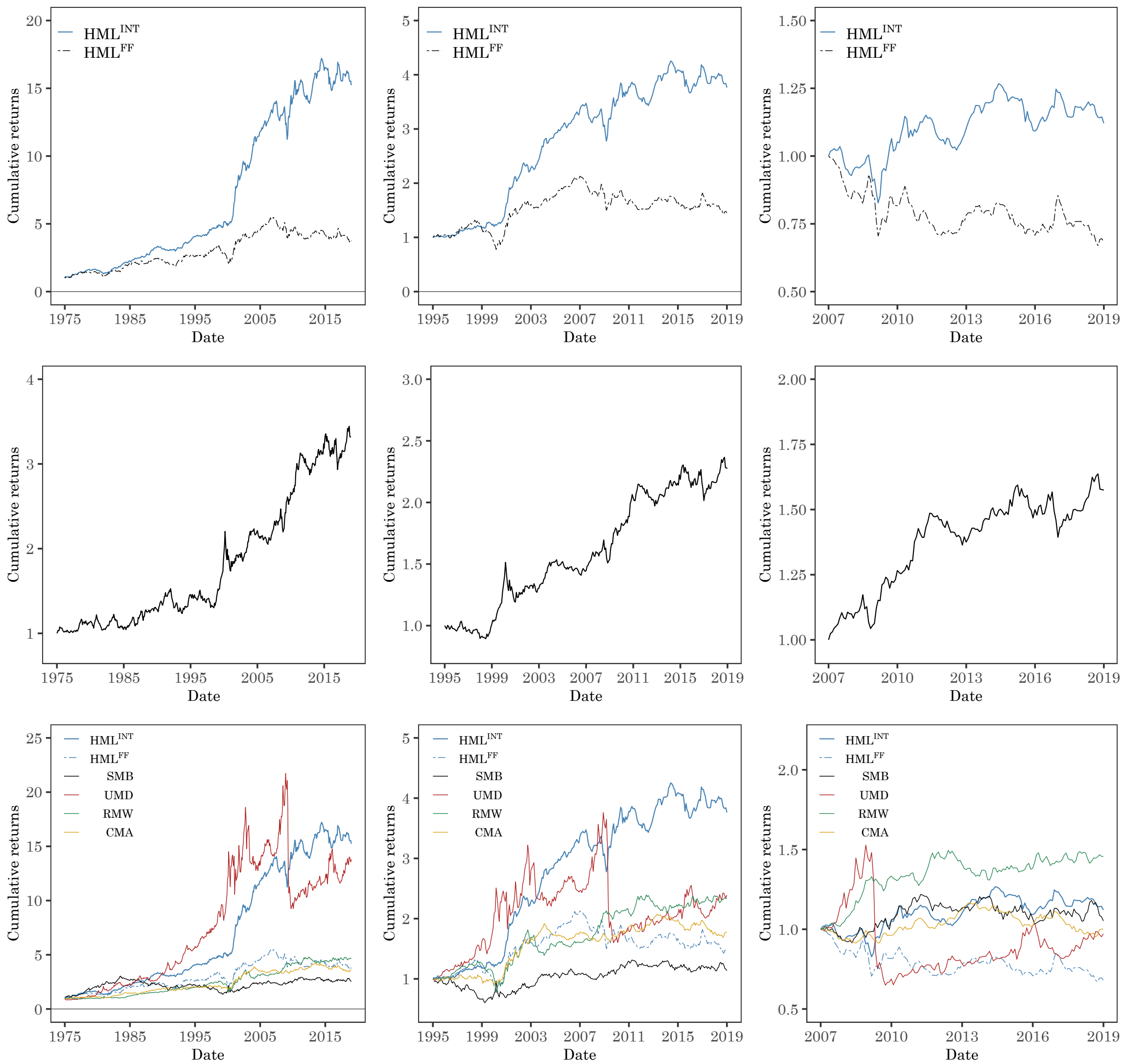


Figure 2.11: Performance of Intangible Value with Industry Filters

Description: This figure plots the performance of HML^{INT} that is formed after dropping financials, utilities, and public service firms from the sample. The top panel plots the cumulative returns of one dollar invested in the HML^{FF} and HML^{INT} portfolios from the beginning of 1975, 1995, and 2007. The middle panel plots the cumulative returns of one dollar invested in the portfolio that is long the HML^{INT} portfolio and short the HML^{FF} portfolio. The bottom panel plots the cumulative returns of one dollar invested in the factors from the three- and five-factor models plus momentum, along with the the HML^{FF} and HML^{INT}.
Interpretation: Intangible value's outperformance is more pronounced when financials, utilities, and public service firms are dropped during portfolio formation.

	(1)	(2)	(3)	(4)
α (%)	12.93	12.56	9.45	9.12
	(4.14)	(3.94)	(3.18)	(3.06)
β_{MktRF}	-0.36	-0.33	-0.11	-0.08
	(-1.14)	(-1.02)	(-0.35)	(-0.26)
β_{SMB}	0.19	0.19	0.23	0.23
	(1.41)	(1.38)	(1.76)	(1.75)
$\beta_{HML^{\text{INDFF}}}$	0.27		0.26	
	(2.71)		(2.60)	
$\beta_{HML^{\text{INT}}}$		0.29		0.30
		(2.87)		(2.88)
β_{UMD}	0.54	0.55	0.54	0.54
	(2.79)	(2.80)	(2.76)	(2.77)
β_{RMW}			0.32	0.32
			(2.83)	(2.90)
β_{CMA}			0.16	0.16
			(1.74)	(1.69)
Adj. R^2	75.38	75.12	79.49	79.84
RMSE	0.41	0.41	0.33	0.33
Prob > χ^2		0.21		0.41

Table 2.D1: Pricing Errors – Industry-Sorted Traditional Value

Description: This table represents pricing results for the Fama and French (1992, 1993, 2015) three factor and five factor models plus momentum. In terms of test assets, columns (1) and (2) use 25 portfolios double-sorted on size and book-to-market and 10 portfolios sorted on momentum. Columns (3) and (4) additionally include 10 investment and 10 profitability portfolios. Fama and MacBeth (1973) T-statistics are reported in parentheses. Prob > χ^2 tests the hypothesis that alphas of the models using either intangible or traditional value factors are significantly different. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve).
Interpretation: Sorting firms within industry improves the asset pricing performance of the traditional value factor.

	Full sample	1975-1994	1995-2006	2007-2018
	(1)	(2)	(3)	(4)
A. $\mathbf{HML_t^{INT}} = \boldsymbol{\alpha} + \boldsymbol{\beta_{HML^{INDFF}}} \cdot \mathbf{HML_t^{INDFF}} + \boldsymbol{\epsilon_t}$				
α (%)	2.16	0.94	4.66	1.83
	(4.89)	(1.43)	(5.35)	(2.32)
$\beta_{HML^{INDFF}}$	0.85	0.89	0.76	0.90
	(33.08)	(27.96)	(13.70)	(23.89)
Adj. R^2	79.27	79.22	77.26	82.30
RMSE	2.97	2.83	3.24	2.77
α/RMSE	0.73	0.33	1.44	0.66
B. $\mathbf{HML_t^{INDFF}} = \boldsymbol{\alpha} + \boldsymbol{\beta_{HML^{INT}}} \cdot \mathbf{HML_t^{INT}} + \boldsymbol{\epsilon_t}$				
α (%)	-1.19	0.42	-3.38	-1.89
	(-2.43)	(0.64)	(-3.18)	(-2.32)
$\beta_{HML^{INT}}$	0.94	0.89	1.01	0.92
	(36.45)	(22.76)	(23.21)	(17.20)
Adj. R^2	79.27	79.22	77.26	82.30
RMSE	3.12	2.84	3.73	2.81
α/RMSE	-0.38	0.15	-0.91	-0.68

Table 2.D2: Single Factor Models – Industry-sorted Traditional Value

Description: In this table, we study the relative performance of the $\mathrm{HML}^{\mathrm{INDFF}}$ and $\mathrm{HML}^{\mathrm{INT}}$ factors. Specifically, we report alphas and betas of a regression of each return on the other, for the full sample as well as for sub-periods around the Internet Bubble and the Great Recession. Firms are sorted within industry first to form the $\mathrm{HML}^{\mathrm{INDFF}}$ factor. The data are monthly and the sample period is 1975 to 2018. We include T-statistics that adjust for heteroskedasticity in parentheses. All factors are annualized in percent per year.
Interpretation: Traditional value factor's returns are marginally improved when employing the within-industry sorting and weighting methodology.

		Full sample	1975-1994	1995-2006	2007-2018
		(1)	(2)	(3)	(4)
HML$^{\mathbf{INDFF}}$	$\mathbb{E}[R]$	4.06	6.08	5.96	-1.20
		(3.93)	(4.37)	(2.64)	(-0.62)
	σ	6.86	6.23	7.82	6.67
	[0.05, 0.95]	[-31.12, 41.26]	[-26.51, 39.72]	[-33.84, 60.48]	[-32.78, 33.41]
	Sharpe	0.59	0.98	0.76	-0.18
HML$^{\mathbf{INT}}$	$\mathbb{E}[R]$	5.60	6.34	9.21	0.76
		(5.70)	(4.57)	(4.70)	(0.40)
	σ	6.52	6.21	6.78	6.57
	[0.05, 0.95]	[-27.54, 40.43]	[-23.63, 40.17]	[-25.95, 48.42]	[-36.38, 35.93]
	Sharpe	0.86	1.02	1.36	0.12
HML$^{\mathbf{IME}}$	$\mathbb{E}[R]$	6.35	7.02	9.30	2.28
		(6.81)	(5.06)	(5.28)	(1.30)
	σ	6.18	6.21	6.10	6.09
	[0.05, 0.95]	[-26.48, 40.80]	[-25.11, 40.98]	[-20.31, 45.42]	[-35.03, 36.87]
	Sharpe	1.03	1.13	1.53	0.37
HML$^{\mathbf{INT}}$ - HML$^{\mathbf{INDFF}}$	$\mathbb{E}[R]$	1.54	0.26	3.25	1.95
		(3.24)	(0.40)	(3.03)	(2.38)
	σ	3.15	2.91	3.72	2.84
	[0.05, 0.95]	[-14.94, 18.00]	[-15.32, 16.37]	[-14.36, 24.97]	[-10.90, 16.72]
	Information	0.49	0.09	0.88	0.69
	Appraisal	0.73	0.33	1.44	0.66
HML$^{\mathbf{IME}}$ - HML$^{\mathbf{INDFF}}$	$\mathbb{E}[R]$	2.29	0.94	3.34	3.48
		(3.37)	(1.10)	(2.10)	(2.73)
	σ	4.50	3.83	5.51	4.41
	[0.05, 0.95]	[-23.18, 26.83]	[-23.26, 22.40]	[-22.67, 37.23]	[-21.28, 26.83]
	Information	0.51	0.25	0.61	0.79
	Appraisal	0.89	0.58	1.40	0.79

Table 2.D3: Performance Statistics – Industry-sorted Traditional Value

Description: This table summarizes the risk and return associated with intangible and traditional value. Firms are sorted within industry first to form the HML$^{\text{INDFF}}$ factor. HML$^{\text{INT}}-$HML$^{\text{INDFF}}$ refers to the portfolio that is long HML$^{\text{INT}}$ and short HML$^{\text{INDFF}}$, and HML$^{\text{IME}}-$HML$^{\text{FF}}$ refers to the portfolio that is long HML$^{\text{IME}}$ and short HML$^{\text{INDFF}}$. The numbers in parentheses are T-statistics for the test that the average return, $\mathbb{E}[R]$, is different from zero. The information ratio is $\mathbb{E}[R_p - R_b]/\sigma(R_p - R_b)$, or the Sharpe Ratio of the long-short portfolio. The appraisal ratio is α/RMSE of a regression of intangible value returns (HML$^{\text{INT}}$ or HML$^{\text{IME}}$) on traditional value returns. The underlying data are monthly and the full sample period is 1975 to 2018. All factors are annualized in percent per year.
Interpretation: Traditional value factor's returns are marginally improved when employing the within-industry sorting and weighting methodology.

	HML^{INDFF}	HML^{INT}	HML^{IME}	HML^{UINT}	HML^{INDFF}	HML^{INT}	HML^{IME}	HML^{UINT}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α (%)	-0.77	2.00	3.17	0.63	-0.97	1.65	2.78	2.06
	(-1.44)	(4.11)	(5.19)	(0.48)	(-1.90)	(3.37)	(4.57)	(1.65)
β_{MktRF}	-0.01	0.00	0.01	0.10	0.01	0.02	0.04	0.07
	(-1.03)	(0.00)	(1.14)	(3.71)	(0.73)	(1.69)	(3.01)	(2.53)
β_{SMB}	-0.04	0.06	0.08	0.35	-0.01	0.08	0.10	0.24
	(-1.89)	(3.38)	(4.18)	(5.83)	(-0.68)	(4.81)	(5.18)	(5.46)
$\beta_{HML^{\mathrm{INT}}}$	0.93				0.83			
	(32.57)				(20.86)			
$\beta_{HML^{\mathrm{INDFF}}}$		0.84	0.69	-0.10		0.76	0.57	-0.03
		(32.16)	(20.74)	(-1.37)		(27.03)	(16.13)	(-0.42)
β_{UMD}	-0.03	0.00	0.01	0.01	-0.03	-0.01	-0.00	0.04
	(-2.02)	(0.31)	(0.57)	(0.41)	(-2.80)	(-0.43)	(-0.19)	(1.21)
β_{RMW}					0.03	0.04	0.02	-0.33
					(0.91)	(1.58)	(0.71)	(-4.78)
β_{CMA}					0.15	0.12	0.19	-0.03
					(4.13)	(4.19)	(4.64)	(-0.39)
Adj. R^2	80.00	79,92	60.81	22.88	81.04	80.82	62.71	28.58
RMSE	3.07	2.92	3.87	8.00	2.99	2.86	3.78	7.70

Table 2.D4: Alphas – Industry-sorted Traditional Value

Description: In this table, we report portfolio alphas and betas of a regression of different variants of HML portfolio returns on traditional factor models. Firms are sorted within industry first to form the HML^{INDFF} factor. Columns (1) through (4) use the Fama and French (1992, 1993) three factor model, and columns (5) through (8) use the Fama and French (2015) five factor model. Both specifications are augmented with the momentum factor. Columns (1) and (5) are benchmarks that set HML^{IME} as the dependent variable and replace the intangibles-adjusted HML factor in the aforementioned models. We include T-statistics that adjust for heteroskedasticity in parentheses. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve).
Interpretation: The alphas for industry-sorted traditional value are negative as in Table 2.11, but are now significant for the five-factor model.

	(1)	(2)	(3)	(4)
α (%)	13.28	12.55	8.59	9.25
	(4.15)	(3.95)	(2.89)	(3.09)
β_{MktRF}	-0.38	-0.33	-0.04	-0.09
	(-1.18)	(-1.03)	(-0.12)	(-0.30)
β_{SMB}	0.18	0.19	0.24	0.23
	(1.36)	(1.40)	(1.78)	(1.75)
$\beta_{HML^{\text{FF}}}$	0.30		0.24	
	(2.35)		(1.92)	
$\beta_{HML^{\text{INT}}}$		0.33		0.33
		(2.82)		(2.73)
β_{UMD}	0.54	0.55	0.53	0.54
	(2.79)	(2.79)	(2.74)	(2.76)
β_{RMW}			0.32	0.32
			(2.87)	(2.88)
β_{CMA}			0.18	0.16
			(1.95)	(1.79)
Adj. R^2	73.14	74.93	78.74	79.46
RMSE	0.43	0.42	0.34	0.33
Prob $> \chi^2$		0.20		0.17

Table 2.D5: Pricing Errors – Excl. Utilities, Financials, and Public Service Firms

Description: This table represents pricing results for the Fama and French (1992, 1993, 2015) three factor and five factor models augmented with the momentum factor. When forming the HML^{INT} portfolio, we drop financial firms (SIC codes 6000-6999), regulated utilities (4900-4999), and firms categorized as public service, international affairs, or non-operating establishments (9000+). In terms of test assets, columns (1) and (2) use 25 portfolios double-sorted on size and book-to-market and 10 portfolios sorted on momentum. Columns (3) and (4) additionally include 10 investment and 10 profitability portfolios. Fama and MacBeth (1973) T-statistics are reported in parentheses. Prob $> \chi^2$ is the p-value of the test that the alpha from the model using HML^{INT} is significantly different from the alpha from the model using HML^{FF}. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve).
Interpretation: Cross-sectional asset pricing performance of intangible value is invariant to dropping nontraditional industries.

	Full sample	1975-1994	1995-2006	2007-2018
	(1)	(2)	(3)	(4)
A. $\mathbf{HML}_t^{\mathbf{INT}} = \boldsymbol{\alpha} + \boldsymbol{\beta}_{\boldsymbol{HML}^{\mathbf{FF}}} \cdot \mathbf{HML}_t^{\mathbf{FF}} + \boldsymbol{\epsilon}_t$				
α (%)	4.66	4.56	7.21	2.85
	(6.29)	(4.58)	(4.87)	(1.82)
$\beta_{HML^{\mathrm{FF}}}$	0.51	0.52	0.46	0.58
	(17.06)	(11.70)	(8.65)	(9.24)
Adj. R^2	51.32	51.01	51.72	50.31
RMSE	4.98	4.60	5.19	5.27
α/RMSE	0.94	0.99	1.39	0.54
B. $\mathbf{HML}_t^{\mathbf{FF}} = \boldsymbol{\alpha} + \boldsymbol{\beta}_{\boldsymbol{HML}^{\mathbf{INT}}} \cdot \mathbf{HML}_t^{\mathbf{INT}} + \boldsymbol{\epsilon}_t$				
α (%)	-2.96	-2.00	-4.86	-3.84
	(-2.74)	(-1.30)	(-1.92)	(-2.07)
$\beta_{HML^{\mathrm{INT}}}$	1.99	0.99	1.14	0.87
	(19.50)	(15.47)	(12.03)	(8.69)
Adj. R^2	51.32	51.01	51.72	50.31
RMSE	6.94	6.36	8.20	6.42
α/RMSE	-0.43	-0.31	-0.59	-0.60

Table 2.D6: Single Factor Models – Excl. Utilities, Financials, and Public Service Firms

Description: In this table, we study the relative performance of the $\mathrm{HML}^{\mathrm{FF}}$ and $\mathrm{HML}^{\mathrm{INT}}$ factors. When forming the $\mathrm{HML}^{\mathrm{INT}}$ portfolio, we drop financial firms (SIC codes 6000-6999), regulated utilities (4900-4999), and firms categorized as public service, international affairs, or non-operating establishments (9000+). We report alphas and betas of a regression of each return on the other, for the full sample as well as for sub-periods around the Internet Bubble and the Great Recession. The data are monthly and the sample period is 1975 to 2018. We include T-statistics that adjust for heteroskedasticity in parentheses. All factors are annualized in percent per year.
Interpretation: An intangible value factor that excludes nontraditional industries exhibits even higher outperformance over the traditional value factor.

		Full sample	1975-1994	1995-2006	2007-2018
		(1)	(2)	(3)	(4)
	$\mathbb{E}[R]$	3.49	5.14	6.99	-2.77
		(2.33)	(2.53)	(2.05)	(-1.05)
HML$^{\text{FF}}$	σ	9.95	9.08	11.80	9.11
	[0.05, 0.95]	[-48.36, 63.24]	[-45.72, 63.12]	[-55.92, 78.24]	[-44.04, 48.84]
	Sharpe	0.35	0.57	0.59	-0.30
	$\mathbb{E}[R]$	6.46	7.22	10.40	1.23
		(6.00)	(4.91)	(4.82)	(0.57)
HML$^{\text{INT}}$	σ	7.14	6.58	7.48	7.48
	[0.05, 0.95]	[-29.78, 41.67]	[-23.24, 41.06]	[-22.13, 53.5]	[-44.31, 39.12]
	Sharpe	0.90	1.10	1.39	0.16
	$\mathbb{E}[R]$	6.68	7.28	9.73	2.64
		(6.88)	(5.19)	(4.90)	(1.49)
HML$^{\text{IME}}$	σ	6.44	6.27	6.87	6.13
	[0.05, 0.95]	[-25.67, 43.40]	[-23.80, 42.22]	[-20.34, 46.91]	[-31.91, 33.58]
	Sharpe	1.04	1.16	1.42	0.43
	$\mathbb{E}[R]$	2.97	2.08	3.41	4.00
		(2.84)	(1.47)	(1.43)	(2.14)
HML$^{\text{INT}}$	σ	6.94	6.35	8.24	6.48
- HML$^{\text{FF}}$	[0.05, 0.95]	[-35.49, 40.01]	[-33.31, 39.27]	[-39.66, 49.29]	[-31.82, 34.57]
	Information	0.43	0.33	0.41	0.62
	Appraisal	0.94	0.99	1.39	0.54
	$\mathbb{E}[R]$	3.19	2.13	2.73	5.40
		(2.78)	(1.39)	(1.05)	(2.58)
HML$^{\text{IME}}$	σ	7.60	6.87	8.98	7.27
- HML$^{\text{FF}}$	[0.05, 0.95]	[-40.89, 44.53]	[-39.84, 38.67]	[-51.31, 53.18]	[-37.27, 40.43]
	Information	0.42	0.31	0.30	0.74
	Appraisal	1.06	1.04	1.35	0.77

Table 2.D7: Performance Statistics – Excl. Utilities, Financials, and Public Service Firms

Description: This table summarizes the risk and return associated with intangible and traditional value. When forming the HML$^{\text{INT}}$ portfolio, we drop financial firms (SIC codes 6000-6999), regulated utilities (4900-4999), and firms categorized as public service, international affairs, or non-operating establishments (9000+). HML$^{\text{INT}}-$HML$^{\text{FF}}$ refers to the portfolio that is long HML$^{\text{INT}}$ and short HML$^{\text{INT}}$, and HML$^{\text{IME}}-$HML$^{\text{FF}}$ refers to the portfolio that is long HML$^{\text{IME}}$ and short HML$^{\text{INT}}$. The numbers in parentheses are T-statistics for the test that the average return, $\mathbb{E}[R]$, is different from zero. The information ratio is $\mathbb{E}[R_p - R_b]/\sigma(R_p - R_b)$, or the Sharpe Ratio of the long-short portfolio. The appraisal ratio is α/RMSE of a regression of intangible value returns (HML$^{\text{INT}}$ or HML$^{\text{IME}}$) on traditional value returns. The underlying data are monthly and the full sample period is 1975 to 2018. All factors are annualized in percent per year.
Interpretation: Intangible value with industry filters outperforms traditional value.

	HML^{FF}	HML^{INT}	HML^{IME}	HML^{UINT}	HML^{FF}	HML^{INT}	HML^{IME}	HML^{UINT}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α (%)	-1.20	3.50	4.24	2.03	-1.56	2.59	3.35	0.80
	(-1.19)	(4.92)	(5.85)	(1.15)	(-1.64)	(3.73)	(4.75)	(0.43)
β_{MktRF}	-0.11	0.05	0.03	0.06	-0.05	0.08	0.07	0.09
	(-5.94)	(3.62)	(2.29)	(1.63)	(-2.48)	(5.57)	(4.33)	(2.30)
β_{SMB}	-0.27	0.21	0.18	0.42	-0.22	0.23	0.20	0.48
	(-8.81)	(10.07)	(8.23)	(7.63)	(-6.67)	(9.48)	(8.40)	(8.28)
$\beta_{HML^{INT}}$	1.04				0.79			
	(22.25)				(12.78)			
$\beta_{HML^{FF}}$		0.58	0.47	0.22		0.46	0.34	0.13
		(22.83)	(16.28)	(3.52)		(14.77)	(10.11)	(1.79)
β_{UMD}	-0.06	-0.01	0.00	-0.05	-0.07	-0.02	-0.02	-0.07
	(-1.95)	(-0.21)	(0.19)	(-0.97)	(-3.06)	(-0.98)	(-0.89)	(-1.43)
β_{RMW}					0.02	0.10	0.08	0.20
					(0.56)	(2.99)	(2.22)	(2.26)
β_{CMA}					0.44	0.25	0.27	0.16
					(7.03)	(5.65)	(5.27)	(1.34)
Adj. R^2	65.34	62.71	51.42	15.23	69.67	65.75	55.44	16.46
RMSE	5.86	4.36	4.49	11.09	5.48	4.18	4.30	11.01

Table 2.D8: Alphas – Excl. Utilities, Financials, and Public Service

Description: In this table, we report portfolio alphas and betas of a regression of different variants of HML portfolio returns on traditional factor models. When forming the HML^{INT} portfolio, we drop financial firms (SIC codes 6000-6999), regulated utilities (4900-4999), and firms categorized as public service, international affairs, or non-operating establishments (9000+). Columns (1) through (4) use the Fama and French (1992, 1993) three factor model, and columns (5) through (8) use the Fama and French (2015) five factor model. Both specifications are augmented with the momentum factor. Columns (1) and (5) are benchmarks that set HML^{IME} as the dependent variable and replace the intangibles-adjusted HML factor in the aforementioned models. We include T-statistics that adjust for heteroskedasticity in parentheses. The sample is monthly from January 1975 to December 2018. All coefficients are reported in percentage per year (monthly percentages multiplied by twelve).
Interpretation: The alphas for HML^{INT} and HML^{IME} with industry filters are positive and significant.

2.E Bibliography

Amenc, Noël, Felix Goltz, and Ben Luyten (2020) "Intangible Capital and the Value Factor: Has Your Value Definition Just Expired?" *The Journal of Portfolio Management*, 46 (7), 83–99.

Ang, Andrew (2014) *Asset management: A systematic approach to factor investing*: Oxford University Press.

Arnott, Robert D, Campbell R Harvey, Vitali Kalesnik, and Juhani T Linnainmaa (2021) "Reports of values death may be greatly exaggerated," *Financial Analysts Journal*, 77 (1), 44–67.

Asness, Clifford S, R Burt Porter, and Ross L Stevens (2000) "Predicting stock returns using industry-relative firm characteristics," Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=213872.

Belo, Frederico, Vito Gala, Juliana Salomao, and Maria Ana Vitorino (2019) "Decomposing firm value,"Technical report, National Bureau of Economic Research.

Bloom, Nicholas and John Van Reenen (2007) "Measuring and explaining management practices across firms and countries," *The Quarterly Journal of Economics*, 122 (4), 1351–1408.

Bound, John, Clint Cummins, Zvi Griliches, Bronwyn H Hall, and Adam B Jaffe (1982) "Who does R&D and who patents?"Technical report, National Bureau of Economic Research.

Corrado, Carol, Charles Hulten, and Daniel Sichel (2005) "Measuring capital and technology: an expanded framework," in *Measuring Capital in the New Economy*, 11–46: University of Chicago Press.

——— (2009) "Intangible capital and US economic growth," *Review of Income and Wealth*, 55 (3), 661–685.

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos (2020) "The cross-section of risk and returns," *The Review of Financial Studies*, 33 (5), 1927–1979.

Daniel, Kent and Sheridan Titman (1997) "Evidence on the Characteristics of Cross Sectional Variation in Stock Returns," *The Journal of Finance*, 52 (1), 1–33.

Eisfeldt, Andrea L. and Dimitris Papanikolaou (2013a) "Internet Appendix to Organization Capital and the Cross-Section of Expected Returns,"Technical report.

——— (2013b) "Organization Capital and the Cross-Section of Expected Returns," *The Journal of Finance*, 68 (4), 1365 – 1406, 10.1111/jofi.12034.

——— (2014) "The value and ownership of intangible capital," *American Economic Review*, 104 (5), 189–94, `http://ideas.repec.org/a/aea/aecrev/v104y2014i5p189-94.html`.

Ewens, Michael, Ryan H Peters, and Sean Wang (2020) "Measuring intangible capital with market prices," Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3287437`.

Falato, Antonio, Dalida Kadyrzhanova, and Jae Sim (2013) "Rising Intangible Capital, Shrinking Debt Capacity, and the US Corporate Savings Glut,"Technical report, FEDS Working Paper No. 2013-67, `http://ssrn.com/abstract=2350863`.

Fama, Eugene F. and Kenneth R. French (1992) "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47 (2), 427–465.

——— (1993) "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33 (1), 3–56.

Fama, Eugene F and Kenneth R French (2015) "A five-factor asset pricing model," *Journal of Financial Economics*, 116 (1), 1–22.

Fama, Eugene F. and James D. MacBeth (1973) "Risk, Return, and Equilibrium: Empirical Tests," *The Journal of Political Economy*, 81 (3), 607–636.

Golubov, Andrey and Theodosia Konstantinidi (2019) "Where Is the Risk in Value? Evidence from a Market-to-Book Decomposition," *The Journal of Finance*, 74 (6), 3135–3186.

Goncalves, Andrei and Gregory Leonard (2020) "The Fundamental-to-Market Ratio and the Value Premium Decline," Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3573444`.

Hou, Kewei, Chen Xue, and Lu Zhang (2015) "Digesting anomalies: An investment approach," *The Review of Financial Studies*, 28 (3), 650–705.

Hulten, Charles R and Xiaohui Hao (2008) "What is a company really worth," *Intangible Capital and the Market to Book Value Puzzles, NBER Working Paper Series*, 14548.

Kogan, Leonid and Dimitris Papanikolaou (2014) "Growth Opportunities, Technology Shocks, and Asset Prices," *The Journal of Finance*, 69 (2), 675–718, https://doi.org/10.1111/jofi.12136.

Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman (2020) "Left Behind: Creative Destruction, Inequality, and the Stock Market," *Journal of Political Economy*, 128 (3), 855–906, 10.1086/704619.

Koh, Ping-Sheng and David M Reeb (2015) "Missing R&D," *Journal of Accounting and Economics*, 60 (1), 73–94.

Lev, Baruch (2000) *Intangibles: Management, measurement, and reporting*: Brookings institution press.

Lev, Baruch and Suresh Radhakrishnan (2005) "The Valuation of Organization Capital," in *Measuring Capital in the New Economy*, 73–110: National Bureau of Economic Research, Inc., `http://ideas.repec.org/h/nbr/nberch/10619.html`.

Lev, Baruch and Anup Srivastava (2019) "Explaining the recent failure of value investing," *NYU Stern School of Business*.

Li, Wendy C. Y. and Bronwyn H. Hall (2020) "Depreciation of Business R&D Capital," *Review of Income and Wealth*, 66 (1), 161–180.

Papanikolaou, Dimitris (2011) "Investment Shocks and Asset Prices," *Journal of Political Economy*, 119 (4), 639–685, 10.1086/662221.

Park, Hyuna (Forthcoming) "An Intangible-adjusted Book-to-market Ratio Still Predicts Stock Returns," *Critical Finance Review*.

Peters, Ryan H and Lucian A Taylor (2017) "Intangible capital and the investment-q relation," *Journal of Financial Economics*, 123 (2), 251–272.

Rampini, Adriano A and S Viswanathan (2013) "Collateral and capital structure," *Journal of Financial Economics*, 109 (2), 466–492.

Rizova, Savina and Namiko Saito (2020) "Intangibles and Expected Stock Returns," Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3697452`.

Solow, Robert M. (1955) "The Production Function and the Theory of Capital," *The Review of Economic Studies*, 103–107.

——— (1957) "Technical Change and the Aggregate Prodution Function," *The Review of Economics and Statistics*, 39 (3), 312–320, `http://www.jstor.org/stable/1926047`.

Zhang, Mindy Xiaolan (2014) "Who Bears Firm-Level Risk? Implications for cash flow volatility," Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2357050.

CHAPTER 3

Deposit Insurance Premiums and Bank Risk

(with Marcelo Rezende)

In the decade following the Great Recession, banking regulation was drastically reformed. Policymakers tightened capital requirements and changed deposit insurance pricing, which led to notable increases in banks' balance sheet costs. These costs narrowed profit margins and may have induced banks to "search for yield" by rebalancing their asset portfolios (Stein, 2013). Whether balance sheet costs affect bank risk has become an important consideration for banking regulation.

However, establishing the causal effect of balance sheet costs on banks' portfolios is difficult because exogenous variation in these costs is scarce. Differences in these costs over time and across banks are correlated with unobservable characteristics that influence banks' behavior. Moreover, measures of balance sheet costs generally depend on bank risk indicators (such as capital and liquidity ratios) that are also correlated with unobservable shocks that affect banks' activities. Estimates of the effects of balance sheet costs that do not account for these correlations, therefore, will most likely be biased.

In this paper, we analyze the impact of one important type of balance sheet cost—deposit insurance premiums charged by the Federal Deposit Insurance Corporation (FDIC)—on banks' portfolio rebalancing behavior. Deposit insurance premiums for banks with total assets less than $10 billion are a linear function of risk measures, and we exploit a kink in this function to estimate the effects of premiums on the composition of banks' liquid assets. Specifically, we use a regression kink design

(RKD) to test whether banks subject to higher premiums attempt to raise their yields by shifting away from highly liquid assets with no risk (reserves) and towards less liquid assets with low risk (interbank loans).

Our results confirm the hypothesis that deposit insurance premiums weaken the demand for reserves and strengthen the supply of interbank loans. We estimate that a 1-basis point increase in the assessment rate charged annually on each dollar of total assets decreases a bank's holdings of excess reserves by about 80 percent (from $5.5 million to $1.1 million), and more than doubles the amount loaned to other banks (from $3.5 million to $8.9 million). These results are economically meaningful and robust to various validation and falsification tests. Taken together, our findings suggest that deposit insurance premiums induce banks to search for yield.

Our paper is related to the growing interest in banks' demand for liquid assets, which is largely motivated by recent episodes of stress in financial markets (Correa et al., 2020; Quarles, 2020; Copeland et al., 2021; d'Avernas and Vandeweyer, 2021). When market dislocations occur, reserves are readily available to meet cash outflows, whereas Treasury securities must be monetized and interbank loans must be repaid before they can be used to settle cash transactions. As a result, banks' demand for reserves increases and their willingness to exchange reserves for other liquid assets such as Treasury securities and interbank loans falls. The recent stress events highlight the importance of understanding what drives banks' allocation of liquid assets, and we provide evidence that balance sheet costs induce substitution between reserves and interbank loans.

This paper also contributes to the literature on the effects of balance sheet costs on bank lending. Heider et al. (2019), Basten and Mariathasan (2020) and Duquerroy et al. (2020) study the behavior of European banks when monetary policy rates dropped below zero, and document that lower monetary policy rates affect bank lending through funding costs. We contribute to this literature with evidence that, under low interest rates, an increase in balance sheet costs unrelated to interest rate changes can also motivate banks to rebalance their asset portfolios. In addition, whereas the existing literature has mainly examined effects on loans to households and businesses, we study effects on banks' composition of liquid assets.[1]

Moreover, our paper is related to the nascent literature on the impact of deposit insurance premiums on bank behavior. Kreicher et al. (2013), Keating and Macchiavelli (2017), Klee et al. (2019), Banegas and Tase (2020), and Kandrac and Schlusche (2021) show evidence that deposit insurance premiums and the characteristics that determine those premiums (such as domestic or foreign ownership) affect banks' demand for liquid assets. Our approach differs because we exploit a kink in the assessment rate schedule to estimate the effects of deposit insurance premiums. This strategy enables us to compare the behavior of similar banks subject to different assessment rates.

Lastly, our paper contributes to the literature on optimal deposit insurance pricing. Optimal deposit insurance premiums are determined by individual bank failure risk (Buser et al., 1981; Kanatas, 1986; Ronn and Verma, 1986; Acharya and Dreyfus, 1989; Chan et al., 1992; Giammarino et al., 1993; Craine, 1995; John et al., 2000; Boyd et al., 2002; Duffie et al., 2003) as well as systemic risks associated with the joint failure of large and systemically important banks (Pennacchi, 2006; Acharya et al., 2010; Allen et al., 2015; Dávila and Goldstein, 2020). To our best knowledge, the literature does not consider the feedback effects of premiums on bank behavior. As bank behavior may affect welfare, our evidence suggests that optimal deposit insurance premiums should also incorporate these effects.

The rest of this paper is organized as follows: Section 3.1 provides an overview of

[1] Recent theoretical articles on post-crisis monetary policy refer to deposit insurance premiums as important examples of balance sheet costs in their models. See Martin et al. (2013), Duffie and Krishnamurthy (2016), Armenter and Lester (2017), Schulhofer-Wohl and Clouse (2018), Afonso et al. (2019), Afonso et al. (2020), and Kim et al. (2020).

deposit insurance premiums, describes how assessment rates are calculated, and discusses how these rates affect the allocation of liquid assets in banks' balance sheets. Section 3.2 summarizes our data and presents summary statistics. Section 3.3 describes our empirical strategy based on the RKD, and Section 3.4 presents our main results as well as robustness tests. Section 3.5 concludes.

3.1 Institutional Background

3.1.1 Deposit Insurance Premiums

Deposit insurance protects depositors from bank runs and failures. In the United States, the FDIC insures deposits at all domestic banks and at some branches and agencies of foreign banks through the Deposit Insurance Fund (DIF), which is maintained by quarterly premiums (called assessments) that insured banks pay. Each bank's assessment is calculated as its assessment base multiplied by an assessment rate. Following the Dodd-Frank Wall Street Reform and Consumer Protection Act (the Dodd-Frank Act), the assessment base of a bank has been defined as the average consolidated total assets minus average tangible equity and some adjustments for banker's banks and custodial banks.[2] Assessment rates are a function of indicators of bank risk, as we describe next.

3.1.2 Assessment Rate Calculation

The assessment rate of a bank with less than $10 billion in total assets is determined by its risk category, which is a function of three capital ratios (total risk-based capital ratio, tier 1 risk-based capital ratio, and leverage ratio) and the CAMELS

[2] From the creation of the FDIC until March 2011, a bank's assessment base was about equal to its total domestic deposits. On April 1, 2011, the current definition of the assessment base was adopted, as required by the Dodd-Frank Act.

composite rating, which summarizes the general condition of a bank.[3] Risk categories range from 1 to 4, with risk category 1 generally containing well-capitalized banks with good CAMELS ratings and risk category 4 generally containing undercapitalized banks with poor CAMELS ratings. The FDIC then assigns each bank an initial base assessment rate, which increases with the risk category of the bank. During our sample period from April 1, 2011, to June 30, 2016, risk category 2, 3, and 4 banks were assessed fixed initial base assessment rates of 14, 23, and 35 basis points per annum, respectively.

For risk category 1 banks, which comprise 81 percent of the banks in our sample, the FDIC first computes an unconstrained initial base assessment rate as a linear function of six risk measures at the bank level.[4] Next, the constrained initial base assessment rate is set at the minimum rate of 5 basis points if the unconstrained initial base assessment rate is below this minimum and at the maximum rate of 9 basis points if the unconstrained initial base assessment rate is above this maximum. As shown by the solid line in Figure 3.1, this rule creates a relationship between the constrained and unconstrained initial base assessment rates that is flat to the left of 5 basis points, increasing with a slope equal to 1 between 5 and 9 basis points, and flat to the right of 9 basis points. Due to the smaller number of observations around the higher kink, we focus our analysis on the lower kink throughout this paper.

After the constrained initial base assessment rate of a bank is calculated, this rate may be adjusted downward for unsecured debt and upward for brokered deposits and debt issued by other institutions. The rate that results from these adjustments is defined as the total base assessment rate. We restrict our sample to banks that are

---

[3] CAMELS ratings are assigned by bank supervisors based on off-site analysis and on-site bank safety and soundness examinations. Supervisors evaluate six main characteristics and assign a rating to each one. The characteristics are Capital adequacy, Asset quality, Management, Earnings, Liquidity, and Sensitivity to market risk, and the respective ratings are called component ratings.

[4] The six measures are different from the ones used to determine the risk category. Further details can be found in Section 3.2.

not subject to any of these adjustments to ensure that the total base assessment rate matches the constrained assessment rate. In other words, the total base assessment rate as a function of the unconstrained initial base assessment rate for banks in our sample is identical to the function in Figure 3.1. Most banks are not subject to any of these adjustments, and our main results hold when we include banks subject to adjustments.[5] In the remaining sections, we refer to the total base assessment rate as the assessment rate and refer to the unconstrained initial base assessment rate as the unconstrained assessment rate.

3.1.3 Assessment Rates, Excess Reserves, and Interbank Lending

Assessment rates raise the cost of assets and narrow profitability, which may affect banks' allocation of liquid assets. Banks can respond to narrower margins by shifting away from safer and more liquid assets with lower returns and towards riskier and less liquid assets with higher returns. We study how assessment rates affect the substitution between excess reserves and loans to other financial institutions in the federal funds market.

A bank with a reserve account at a Federal Reserve Bank can deposit funds and earn interest on excess reserves (IOER). A bank without a reserve account can only receive IOER via accounts managed by a correspondent bank, which charges a fee for this service. Alternatively, a bank can lend funds overnight to another financial institution in the federal funds market and earn the interest negotiated between the two parties. Banks may earn a higher rate lending to other banks than they would earn holding reserves in their own accounts or through correspondent

[5] The unsecured debt adjustment (UDA) is the only one of these three adjustments that may affect our estimates, because the UDA attenuates the changes in slope of the initial base assessment rate as a function of the unconstrained initial base assessment rate. In contrast, the brokered deposit adjustment (BDA) only applies to banks in risk categories 2 to 4, and the depository institution debt adjustment (DIDA) does not depend on the initial base assessment rate. 3.C describes the calculation of assessment rates in more detail.

banks.[6] More generally, banks demand a premium when lending funds to other financial institutions because those borrowers might default on their loans, whereas reserves held at Federal Reserve Banks are considered free of credit risk. Reserves can also be used immediately to meet cash outflows, whereas interbank loans must be repaid before they can be used to settle transactions.

Due to these institutional differences, the choice between excess reserves and interbank loans can be interpreted as a trade-off between the lower risk and higher liquidity of the former and the larger return of the latter. Deposit insurance premiums affect this trade-off because the assessment base of a bank increases with its total assets: when the assessment rate of a bank rises, the average cost of each dollar of assets also increases.

3.2 Data

The unit of observation in our panel data is a bank-quarter pair. We study the period between the second quarter of 2011 and the second quarter of 2016, because the relevant rule for calculating assessment rates was introduced on April 1, 2011, and revised again on July 1, 2016. Therefore, our estimates are not driven by changes in regulation, which may be correlated with unobservable characteristics of banks or the economy that might also affect banks' behavior.

We additionally restrict our sample based on four bank characteristics. First, we limit the sample to domestic commercial banks to ensure that all institutions in our sample are subject to a homogeneous regulatory framework and that we can observe data on their relevant characteristics. Second, we limit our sample to risk category 1 banks because, as explained in Section 3.1.2, this is the only risk category for which

[6] See 3.B for a discussion of the relationship between IOER rate and the effective federal funds rate (EFFR), a measure of interest rates for interbank loans.

assessment rates vary across banks. Third, we eliminate newly insured institutions from the sample, which are defined as banks that became insured within the past five years at the time of calculation. Newly insured banks are uniformly assigned an assessment rate of 9 basis points if they are in risk category 1.

Fourth, we limit the sample to banks with total assets between \$100 million and \$5 billion. We drop banks with less than \$100 million in assets because a large majority of those banks do not have reserve accounts at Federal Reserve Banks and thus cannot hold reserves with the Federal Reserve.[7] We eliminate banks with more than \$5 billion in assets to ensure that all banks in the sample follow the same schedule of assessment rates. As discussed in Section 3.1.2, banks with more than \$10 billion in total assets, which the FDIC defines as large and highly complex institutions, must follow a schedule that uses bank data that are not readily available from regulatory filings. We drop banks with assets between \$5 billion and \$10 billion because they may choose the schedule for large and highly complex institutions under certain conditions. This restriction only causes a modest decrease in our final sample, because less than 4 percent of commercial banks held more than \$5 billion in total assets during our sample period.

For each bank-quarter observation, we calculate the corresponding unconstrained assessment rate following the FDIC's rule.[8] Figure 3.2 plots the number of bank-quarter observations around thresholds for the minimum and maximum assessment rates (5 and 9 basis points, respectively), using 0.33-basis point bins with an average size of 716 observations. Importantly, the distribution is heavily skewed and the number of observations around the 5-basis point threshold is much higher than the

---

[7] Only 28 percent of observations from banks with less than \$100 million in total assets are from banks with reserve accounts, whereas 59 percent of the observations from banks with assets between \$100 million and \$5 billion are from banks with reserve accounts. Of note, our results are weaker when we include banks with less than \$100 million of total assets in the sample.

[8] We discuss these calculations in Appendix 3.C.3.

number around the 9-basis point threshold. Because the RKD estimation strategy requires a large number of observations around treatment thresholds, we restrict our analysis to the 5-basis point threshold.

As shown in Table 3.7, the unconstrained assessment rate is calculated using seven risk measures: tier 1 leverage ratio, ratio of loans past due 30-89 days to gross assets, ratio of nonperforming assets to gross assets, adjusted brokered deposits ratio, ratio of net loan charge-offs to gross assets, ratio of net income before taxes to risk-weighted assets, and the weighted average CAMELS component rating. The adjusted brokered deposits ratio is given by total brokered deposits divided by total deposits and is adjusted for four-year cumulative total gross asset growth. Additionally, the ratio of net loan charge-offs to gross assets and the ratio of net income before taxes to risk-weighted assets are adjusted for mergers that occurred during the measurement period and incorporate charge-off and income flows for the trailing four quarters. All financial ratios for the current quarter, except the four-quarter trailing flows, are computed using balance sheet data from the end of the previous quarter contained in the Consolidated Reports of Condition and Income (FFIEC 031 and 041), also known as the Call Reports.

Following the FDIC's rule, we take each bank's most recent CAMELS composite rating from confidential Federal Reserve data to determine its risk category. Then, for each risk category 1 bank, we take the most recent weighted average CAMELS component rating to compute the unconstrained assessment rate for a given quarter. The weighted average CAMELS component rating is calculated using weights outlined in Table 3.8.

Our dependent variables measure reserve holdings and interbank lending activity in the federal funds market. We use confidential data on the dollar amounts of reserves and excess reserves held by banks with the Federal Reserve in the last week of each quarter, as well as the average amounts of reserves and excess reserves each quarter.[9] We collect quarterly data on the amounts of federal funds sold and purchased, of securities purchased under agreements to resell (repo), and of securities sold under agreements to repurchase (reverse repo) from the Call Reports.

We also use the total capital ratio and the tier 1 capital ratio to measure bank capitalization, and return on assets (ROA) and return on equity (ROE) to measure profitability. We build these measures with Call Report data to investigate whether variations in the average bank characteristics are smooth around the 5-basis point cutoff.

Table 3.1 summarizes the data. The mean unconstrained assessment rate, equal to 6.05 basis points, is close to the 5-basis point threshold, which is expected given the large number of observations close to this threshold. The means of capital ratios and profitability measures are high and the means of net charge-offs and nonperforming loans ratios are low, consistent with the fact that, on balance, risk category 1 banks are the most capitalized, profitable, and safe. Of note, the number of observations for our various measures of reserves is less than 20,000, whereas the number of total observations in the sample exceeds 30,000, consistent with the fact that about 60 percent of the banks in our sample have reserve accounts at Federal Reserve Banks.

3.3 Regression Kink Design

3.3.1 RKD Estimator

We estimate the effects of deposit insurance premiums using a sharp RKD, as opposed to a fuzzy RKD, because assessment rates are assigned deterministically based on the

[9] A bank's excess reserves is, for the most part, equal to its average end-of-day account balances due from Federal Reserve Banks less its reserve balance requirement (RBR). Balance data are from internal Federal Reserve accounting records whereas bank-level RBR is calculated based on confidential filings of the FR 2900 Report of Transaction Accounts, Vault Cash and Other Deposits.

method described in Sections 3.1 and 3.2. In particular, we use the RKD estimator from Calonico et al. (2014). For each bank i and quarter t, with $i = 1, 2, ..., I$ and $t = 1, 2, ..., T$, X_{it} is the unconstrained assessment rate—the score, forcing, assignment, or running variable in our setting—such that the bank-quarter pair $\{i, t\}$ is subject to the minimum rate of 5 basis points if $X_{it} < 5$ and X_{it} basis points if $X_{it} \geq 5$. In simpler terms, the rate schedule implies that unconstrained assessment rates lower than 5 basis points will be fixed at the floor rate of 5 basis points. We further define:

$$\mu(x) \equiv \mathbb{E}[Y_{it}|X_{it} = x], \tag{3.1}$$

$$\mu_+^{(\nu)} \equiv \lim_{x \to 5^+} d^\nu \mu(x)/dx^\nu, \tag{3.2}$$

$$\mu_-^{(\nu)} \equiv \lim_{x \to 5^-} d^\nu \mu(x)/dx^\nu. \tag{3.3}$$

As described in Card et al. (2015) and Landais (2015), the denominator of the RKD estimand is deterministic; it is the change in the slope of the schedule at the kink, which is equal to 1 at the 5-basis point threshold. Thus, we only need to estimate the numerator of the estimand, namely the change in the slope of the conditional expectation function $\mu(x)$ at the kink, $\tau \equiv \mu_+ - \mu_-$. The bias-corrected local quadratic estimator is as follows:

$$\hat{\tau}(h_{IT}) \equiv \hat{\mu}_{+,2}^{(1)}(h_{IT}) - \hat{\mu}_{-,2}^{(1)}(h_{IT}) - h_{IT}^2 \hat{B}(h_{IT}, b_{IT}), \tag{3.4}$$

where $\hat{\mu}_{+,2}^{(1)}(h_{IT})$ and $\hat{\mu}_{-,2}^{(1)}(h_{IT})$ are local-quadratic estimators of $\mu_+^{(1)}$ and $\mu_-^{(1)}$, respectively, and h_{IT} is a positive bandwidth. $h_{IT}^2 \hat{B}(h_{IT}, b_{IT})$ is a term intended to correct the bias in the estimator caused by the mean-squared-error optimal choice of the bandwidth for $\hat{\mu}_{+,2}^{(1)}(h_{IT}) - \hat{\mu}_{-,2}^{(1)}(h_{IT})$. $\hat{B}(h_{IT}, b_{IT})$ is given by:

$$\hat{B}(h_{IT}, b_{IT}) \equiv \hat{\mu}_{+,3}^{(3)}(b_{IT}) \mathcal{B}_+(h_{IT})/3! - \hat{\mu}_{-,3}^{(3)}(b_{IT}) \mathcal{B}_-(h_{IT})/3! \tag{3.5}$$

where b_{IT} is a pilot bandwidth, $\hat{\mu}^{(3)}_{+,3}(b_{IT})$ and $\hat{\mu}^{(3)}_{-,3}(b_{IT})$ are the local-cubic estimators of $\mu^{(3)}_{+}$ and $\mu^{(3)}_{-}$, respectively, and $\mathcal{B}_{+}(h_{IT})$ and $\mathcal{B}_{-}(h_{IT})$ are asymptotically bounded observed quantities.[10] We estimate τ using local linear and quadratic estimators, clustering standard errors at the bank level, and using the software packages described in Calonico et al. (2017).[11]

3.3.2 Smoothness Assumption of the RKD

The key identifying assumption of the sharp RKD is that the density of the running variable conditional on the unobservable determinants of the outcome variable is sufficiently smooth—that is, continuously differentiable—at the cutoff (Card et al., 2015). This smoothness condition is violated if the density of the running variable has a kink or a discontinuity at the cutoff. Such violation would suggest that individuals can precisely manipulate the running variable at the cutoff. In our context, this assumption requires that banks cannot lower their unconstrained assessment rates in a neighborhood of the 5-basis point threshold.

Figure 3.2 shows that the distribution of the running variable is smooth around 5 basis points, indicating that banks do not manipulate their unconstrained assessment rates within a narrow neighborhood of this threshold. This finding is expected because the unconstrained assessment rate is determined by variables that depend on market prices and decisions of bank supervisors and borrowers, making it difficult for banks to adjust these variables with precision. We also formally test whether the density of unconstrained assessment rates is continuous around this threshold. These tests, which we present in Table 3.9 in 3.D, do not reject the null hypothesis that the density of the running variable is continuous at the threshold of 5 basis points, offering

[10] These quantities are defined in Lemma A.1(B) of Calonico et al. (2014).

[11] The description of the local linear estimator is analogous to description of the local quadratic estimator, and we omit it from the paper for the sake of brevity.

additional support to our RKD.

The smoothness assumption also implies that the expectation of any variable that should not be affected by treatment conditional on the running variable must be twice continuously differentiable at the cutoff. Figures 3.3 and 3.4 examine this hypothesis graphically showing the mean values of covariates as a function of the running variable. In Figure 3.3, we analyze the risk measures that determine the value of the running variable, except the adjusted brokered deposit rate, which is equal to 0 for all banks in our sample. In Figure 3.4, we examine the total capital ratio and the tier 1 capital ratio, which, together with the tier 1 leverage ratio, determine each bank's capital group and risk category. The figures show the mean values of these covariates in the year-quarter that the running variable is measured. Figure 3.4 also includes two measures of profitability, namely ROA and ROE, measured in the previous year-quarter. We use lagged values for ROA and ROE because, in principle, deposit insurance premiums could lower profitability, even though these effects should be modest as assessment rates are small compared to the means of both ratios

The two figures show that the relationships between the running variable and the conditional expectations of those covariates are smooth around the cutoff. In 3.D, we formally test whether these conditional expectations are twice continuously differentiable around the threshold of 5 basis points by estimating treatment effects on those covariates using the estimator $\hat{\tau}(h_{IT})$ and the cutoff of 5 basis points. As shown in Table 3.10, the tests for all covariates do not reject the null hypothesis of no treatment effects, providing further support to our RKD.

3.4 Results

In this section, we examine the effects of assessment rates on banks' excess reserves and interbank lending. We measure the running variable (unconstrained assessment rate) in quarter t and the outcome variables (excess reserves and federal funds sold and purchased) in quarter $t+1$. Our empirical strategy assumes that banks consider their assessment rates in quarter t to be reliable approximations of their rates in $t+1$, which in turn apply to their assessment bases in $t+1$ and affect their decisions about reserve amounts and interbank lending in $t+1$. This assumption is adequate as unconstrained assessment rates are highly correlated over time at the bank level (see 3.E).

3.4.1 Effects of Assessment Rates on Banks' Excess Reserves

We first provide graphical evidence that assessment rates affect banks' excess reserves. Figure 3.5 shows the relationship between the unconstrained assessment rate in t and the natural logarithm of quarter-end excess reserves in t+1. Unconstrained assessment rates, shown in the horizontal axis, are divided into 30 0.67-basis point wide buckets. The mean value of the natural logarithm of excess reserves in each bucket, as well as 95-percent confidence intervals, are shown in the vertical axis.

The figure shows that excess reserves change little on average with unconstrained assessment rates when the constrained rates are constant (left of the 5-bps cutoff). Additionally, excess reserves decrease with unconstrained assessment rates when assessment rates are not constrained by the 5-basis point minimum (right of the 5-basis point cutoff). The decrease in the slope of the relationship between assessment rates and excess reserves at the 5-basis point threshold indicates that assessment rates weaken banks' demand for reserves. This supports the hypothesis that balance sheet costs induce banks to search for yield. Still, the figure shows large dispersion in excess reserve amounts relative to the change in slope, and we leave a more definitive conclusion to the regression analysis.

Table 3.2 shows estimates of the effects of assessment rates on banks' demand for reserves. Following Card et al. (2015) and Landais (2015) among others, we present results with linear and quadratic polynomials to evaluate whether the estimates depend on assumptions about the order of the polynomial. Columns 1 and 2 show estimates using local linear polynomials ($p = 1$) and columns 3 and 4 show estimates using local quadratic polynomials ($p = 2$). In columns 1 and 3, the dependent variable is the natural logarithm of quarter-end excess reserves, and in columns 3 and 4, the dependent variable is the natural logarithm of quarterly average excess reserves.

Estimates of the effects of assessment rates on excess reserve amounts are large, statistically significant, and have the expected sign. The -1.579 estimate of τ in column 1 implies that a 1-basis point increase in the assessment rate decreases the excess reserves of the average bank in the sample from \$5.5 million to \$1.1 million. The coefficient estimate in column 2, equal to -1.694, implies an effect of roughly the same size. The point estimate in column 3, equal to -2.680, implies that a 1-basis point increase in the assessment rate lowers the excess reserves of the average bank in our sample from \$5.5 million to \$0.4 million. The coefficient when using average excess reserves is similar at -2.814. Moreover, we observe similar confidence intervals between columns 1 and 2 and between columns 3 and 4.

Although these changes seem large given the level of banks' excess reserve holdings, they are reasonable considering the distributions of assessment rates and excess reserves. A 1-basis point increase in the assessment rate is roughly one-half of a standard deviation of unconstrained assessment rates in our sample, and the implied decrease in excess reserves (ranging from \$4.4 million to \$5.1 million) is less than 15 percent of a standard deviation of excess reserves. In other words, we can interpret the results as a standard deviation increase in assessment rates lowering excess reserves by one-third of a standard deviation.

The estimates in Table 3.2 indicate that using quarter-end or quarterly average amounts yield similar results, consistent with the fact that quarter-end effects on reserve balances are generally modest at small banks. Because the results using the two alternative measures of reserves are similar, we henceforth only present results using quarter-end measures.

The coefficient estimates in Table 3.2 also show that our results are robust to changes in the order of the polynomial employed. However, the effects implied by the estimates in columns 1 and 2 are smaller than the effects implied by the estimates in columns 3 and 4. Because the estimates with a local linear polynomial appear to be more conservative under our setting, we mostly focus on this specification in the remainder of the paper.

3.4.2 Effects of Assessment Rates on Interbank Lending

We next examine whether assessment rates affect short-term lending in the federal funds market. Figures 3.6 and 3.7 present the relationship between the unconstrained assessment rate and the natural logarithms of federal funds sold and purchased, respectively. Figure 3.6 shows that federal funds sold decrease with the unconstrained assessment rate when the constrained rates are constant (left of the 5-basis point cutoff), suggesting that bank risk weakens interbank loan supply when assessment rates do not vary with bank risk. This negative relationship is consistent with the evidence from Figure 3.5 that excess reserves increase with unconstrained rates on the left of the cutoff, and implies that riskier banks reduce their supply of interbank loans to hold more excess reserves for precautionary reasons.

Figure 3.6 also shows that federal funds sold rise with the unconstrained assessment rate when assessment rates are not constrained by the 5-basis point minimum (right of the 5-basis point cutoff). Once again, this positive relationship is consistent with the negative relationship between excess reserves and assessment rates on the right of the cutoff shown in Figure 3.5, and indicates that deposit insurance premiums motivate banks to take on more risk by shifting their allocations of liquid assets from excess reserves to interbank loans.

Meanwhile, federal funds purchased do not seem to respond to assessment rates, as shown in Figure 3.7. Although the weak relationship between interbank borrowing and assessment rates contrasts with the strong relationship between interbank lending and rates seen in Figure 3.5, the result can be explained by the fact that, on average, small banks sell more funds than they buy.

Table 3.3 shows estimates of the effects of assessment rates on interbank lending. In Panel A, columns 1 and 3 use the natural logarithm of federal funds sold and columns 2 and 4 use the natural logarithm of federal funds purchased as the dependent variable. Consistent with Figure 3.6, estimates of τ using federal funds sold as the dependent variable are positive and statistically significant. The local linear estimate in column 1, equal to 0.949, indicates that the amount of federal funds sold at the average bank would jump from \$3.5 million to \$8.9 million following a 1-basis point increase in its assessment rate. The local quadratic estimate in column 3, equal to 1.212, implies that the federal funds sold would jump to \$11.6 million in response to a 1-basis point increase in assessment rates.

Analogous to our interpretation in Section 3.4.1, these increases in federal funds sold are large given the average amount of federal funds sold, but not relative to the distributions of those two variables. A 1-basis point change in the assessment rate is about one-half of a standard deviation of the unconstrained assessment rates in our sample. The increase in federal funds sold caused by this change in assessment rates (ranging from \$5.4 million to \$8.1 million) is roughly equal to one-third or one-half of a standard deviation of federal funds sold by small banks. Thus, these results imply that a standard deviation increment in assessment rates raises the amount of federal funds sold by one-third to one-half of a standard deviation.

Consistent with Figure 3.7, estimates of τ using the federal funds purchased as the dependent variable, in columns 2 and 4 of Panel A, are closer to zero and not statistically significant. These estimates indicate that the amount of federal funds purchased by the banks in our sample does not respond to assessment rates. In Panel B, we repeat the regressions from Panel A adding reverse repo and repo amounts to federal funds sold and purchased, respectively. Reverse repos and repos are alternative operations that banks conduct to lend and borrow short-term funds. We include reverse repos and repos in the dependent variables to examine whether the estimates in Panel A change if we consider a broader set of short-term borrowing and lending operations. The coefficient estimates in Panel A and B are very close, which can be attributed to the fact that small banks rarely engage in repo operations. Overall, the findings in this section are consistent with the fact the banks in our sample—generally small, safe and sound banks—are much more likely to sell federal funds than to purchase them. In addition, the findings are consistent with the evidence that small banks typically do not purchase federal funds, as discussed in Keating and Macchiavelli (2017) and others.

3.4.3 Validation and Falsification Tests

In this section, we present three additional validity tests for our RKD methodology: estimating treatment effects with placebo cutoffs, using different bandwidth choice procedures, and excluding observations near the cutoff.[12] Similar to the smoothness assumption tests from Section 3.3.2, these three tests support the assumptions of our RKD.

3.4.3.1 Placebo Cutoffs

We first examine whether our estimates of treatment effects are significant at false (or "placebo") cutoff values. Estimates with placebo cutoffs help evaluate whether

[12] Together with the tests of continuity of the score variable and of null treatment effects on pre-treatment and placebo outcomes in Section 3.3.2 and 3.D, these tests constitute the five validation and falsification tests that Cattaneo et al. (2018) introduce for regression discontinuity and RKD designs.

the RKD assumption of continuity of regression functions for treatment and control observations at the cutoff in the absence of treatment hold. Even though this assumption cannot be tested directly, evidence of discontinuities would indicate that it does not hold. Conversely, evidence of continuity away from the cutoff, which is neither necessary nor sufficient for continuity at the cutoff, would offer some support to that assumption.

We examine continuity away from the cutoff by estimating the effects of assessment rates after replacing the true cutoff of 5 basis points with values at which no treatment occurs. We present results using excess reserve holdings and the amounts of federal funds sold as dependent variables because the results in Table 3.3 do not indicate that assessment rates affect the amounts of federal funds purchased. Additionally, we only show estimates with a local linear polynomial because the conclusions using a local quadratic polynomial are about the same.

Figure 3.8 shows our estimates using alternative cutoffs from 2 basis points to 8 basis points. The left and right panels use the natural logarithms of excess reserves and federal funds sold as dependent variables, respectively. Red dots show our point estimates of treatment effects and the vertical lines show robust 95 percent confidence intervals. We report complete regression results in Table 3.11 of Appendix 3.D.2.

In both panels, the confidence intervals do not include zero only when we use the true cutoff of 5 basis points, which supports the assumption of continuity. Of note, the larger number of observations in our data closer to that cutoff, as shown in Figure 3.2, helps to narrow confidence intervals and reject the hypothesis of no treatment effect in that neighborhood. Still, the evidence from this figure favors the assumption of continuity.

3.4.3.2 Sensitivity to Bandwidth Choice

We next examine whether our results are robust to changes in the procedure used to select bandwidths. Different procedures can affect results by generating bandwidths of different lengths: a widening in bandwidths increases the bias of the local polynomial estimator and lowers the variance of the estimator. Accordingly, a widening in bandwidths generally narrows and displaces confidence intervals.

We compare the results using four alternative procedures: one common coverage error (1CCER)-optimal bandwidth selector, one common mean squared error (1CMSE)-optimal bandwidth selector (also used in Tables 3.2 and 3.3), two different coverage error (2DCER)-optimal bandwidth selectors, and two different mean squared error (2DMSE)-optimal bandwidth selectors. As discussed in Cattaneo et al. (2018), mean squared error (MSE)-optimal bandwidth selectors have highly desirable properties for point estimation of treatment effects, but they also have serious disadvantages for building confidence intervals, whereas coverage error (CER)-optimal bandwidth selectors yield point estimators with too much variability relative to their biases, but also generate confidence intervals with better properties than the MSE-optimal bandwidth selectors. Meanwhile, results using one common bandwidth and two different bandwidths may vary meaningfully. For these reasons, we present results using the four possible combinations of MSE-optimal versus CER-optimal bandwidth and one common bandwidth versus two different bandwidth selectors.

Table 3.4 shows the results using the alternative bandwidth selection procedures. Panels A and B present estimates using the natural logarithms of excess reserves and federal funds sold as the dependent variables, respectively. In both panels, the bandwidths are longer and the point estimates of the RKD effect are closer to 0 when we use an MSE-optimal procedure—the better procedure for point estimation—instead of a CER-optimal procedure. Importantly, the point estimates in Panel A remain negative and large in absolute terms and the point estimates in Panel B remain positive and large in absolute terms across the four different procedures.

The confidence intervals in Table 3.4 indicate that our findings from Tables 3.2 and 3.3 are robust to changes in the bandwidth selection procedures. The confidence intervals exclude zero in all procedures except when we estimate the RKD effect on excess reserves using two different MSE-optimal bandwidths. However, as Cattaneo et al. (2018) discuss, the CER-optimal bandwidth is more appropriate than the MSE-optimal bandwidth for validation and falsification purposes because our objective is to test the null hypothesis of no effect and point estimates are less important. Thus, we conclude that our results from Tables 3.2 and 3.3 are also robust to changes in the bandwidth selection procedure.

3.4.3.3 Sensitivity to Observations near the Cutoff

Our last robustness test investigates whether the estimates of the effects of assessment rates on reserves and federal funds sold change materially if we drop observations very close to the 5-basis point cutoff. This exercise, often known as the donut hole approach, tests whether systematic manipulation of assessment rates by banks drives our results. The test assumes that observations close to the cutoff are more likely to be of banks that manipulated their unconstrained assessment rates. Removing observations close to the cutoff would therefore drop observations that are more likely subject to manipulation. Again, manipulation should be a minor concern in our setting because assessment rates are determined by many variables that banks cannot control precisely (e.g. fraction of loans past due and supervisory ratings). However, such exercise helps assess the sensitivity of the results to the extrapolation inherent to local polynomial estimation, in which the observations close to the cutoff influence the results substantially.

Figure 3.9 shows how point estimates and confidence intervals change as we drop observations up to 0.020 basis points on either side of the 5-basis point cutoff. The top and bottom rows show results using MSE-optimal and CER-optimal bandwidths, respectively. The RKD estimate using MSE-optimal bandwidths and excess reserves as the dependent variable is somewhat sensitive to the removal of observations near the cutoff, as the confidence intervals include zero even when we drop observations within a 0.005-basis point radius around the cutoff (top-left panel).

As discussed in Section 3.4.3.2, confidence intervals constructed with CER-optimal bandwidths have better properties than MSE-optimal intervals. The bottom row panels provide evidence that our estimates remain statistically significant even if drop observations from a wider interval around the 5-basis point cutoff. We find that dropping observations up to a 0.015-basis point radius does not change the signs of the estimates on excess reserves and federal funds sold or induce the robust confidence intervals to include zero. In summary, the results from Table 3.3 are mostly unchanged when we remove observations close to the 5-basis point cutoff.

3.4.4 Discussion

In this section, we discuss how our findings relate to optimal deposit insurance systems and pricing. Our results shed some light on this topic, but a comprehensive framework that accounts for the benefits of deposit insurance premiums would be more appropriate to determine what assessment rates and systems would maximize social welfare.

Optimal deposit insurance premiums should depend on the effects of bank behavior on welfare, which we don't consider in this paper. For example, optimal premiums should also account for the potential benefits of interbank loans. In fact, a higher supply of interbank loans helps keep the federal funds market liquid, allowing banks to meet reserve requirements and avoid costly government interventions.

Of note, our estimates on the effects of deposit insurance premiums on bank behavior likely depend on the characteristics of banks around the 5-basis point threshold.

The FDIC sets its assessment rate schedule—including the threshold—to make premiums actuarially fair, that is, to make the DIF break even in expectation.[13] However, an assessment rate schedule that is actuarially fair potentially differs from a schedule that maximizes welfare. This difference implies that our estimates might not hold in an environment in which assessment rates are set to maximize welfare.[14]

In addition, a more comprehensive framework is necessary to design an optimal deposit insurance system. We do not consider several benefits of deposit insurance in this paper. For example, deposit insurance helps prevent bank runs by enabling banks to liquidate assets in an orderly manner. The effect of premiums that we examine thus constitutes one component of a cost and benefit analysis of deposit insurance. In fact, several prominent models that study these costs and benefits abstract from premiums and assume that the government provides deposit insurance for free (Keeley, 1990) or funds it through taxes levied on depositors (Diamond and Dybvig, 1983).

3.5 Conclusion

This paper examines the impact of deposit insurance premiums on banks' demand for reserves and interbank lending in the federal funds market. By exploiting a kink in the schedule of deposit insurance assessment rates, we show that these premiums reduce the demand for reserves and increase the supply of federal funds. The economic significance of our results are large, indicating that balance sheet costs can induce banks to search for yield. Given that larger banks—those outside the scope of this paper—generally have much higher reserve balances and participate more ac-

[13] The schedule of assessment rates over the sample period was largely affected by the Dodd-Frank Act, which required the FDIC to increase rates in order to restore the DIF and to have the cost of this transition be borne by large banks. More generally, the DIF has historically been funded only by assessment fees from banks.

[14] Chan et al. (1992), Craine (1995), and Dávila and Goldstein (2020) discuss differences between optimal deposit insurance premiums and actuarially fair premiums.

tively in the federal funds market, the impact of deposit insurance premiums may be economically substantial.

Meanwhile, the effects of deposit insurance premiums on bank behavior are likely stronger in a low-interest rate environment. Over the period we study, low interest rates kept banks' net interest margins narrow, implying that changes in assessment rates of a few basis points could drive material changes in bank behavior. It is plausible that these effects would be weaker under higher interest rates.

Our findings have an important policy implication. We show that small changes in deposit insurance premiums can meaningfully reduce banks' demand for reserves, a highly liquid asset with negligible credit risk, and raise demand for interbank loans, a less liquid asset with credit risk. Because balance sheet costs imposed by deposit insurance premiums can induce banks to search for yield, optimal deposit insurance pricing should account for the feedback effects of deposit insurance premiums on bank risk.

3.A Figures and Tables

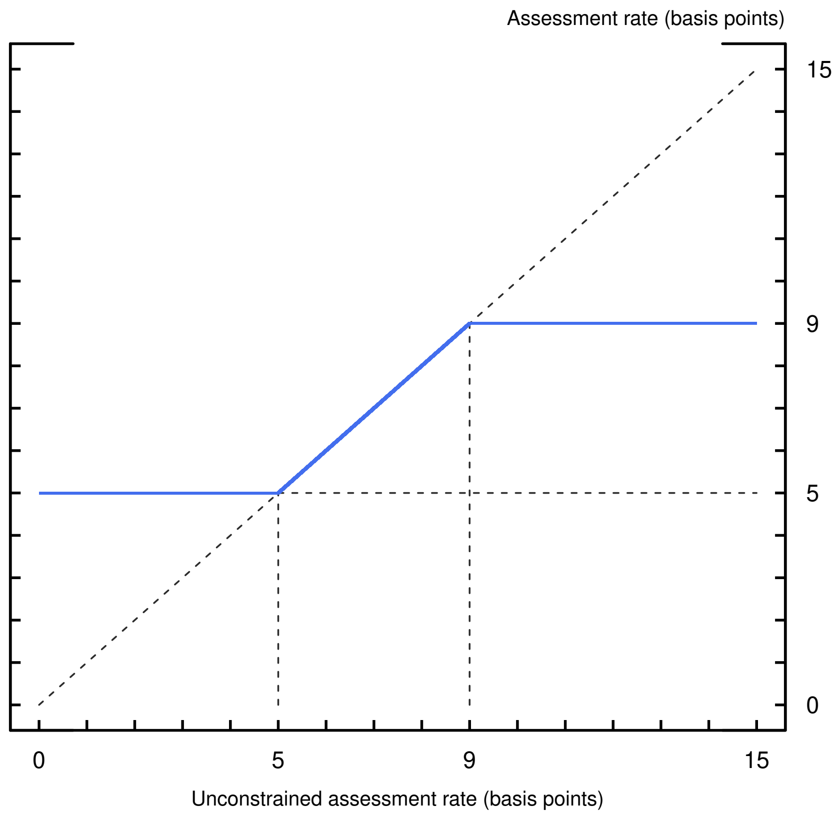


Figure 3.1: Kinks in Deposit Insurance Assessment Rate Schedule

NOTE: The solid line shows the assessment rate as a function of the unconstrained assessment rate for insured risk category 1 banks between April 1, 2011, and June 30, 2016, with total assets below $10 billion. Newly insured institutions (those that became insured within five years) are subject to different rates and are not included in the analysis.
SOURCE: Federal Deposit Insurance Corporation (2011).

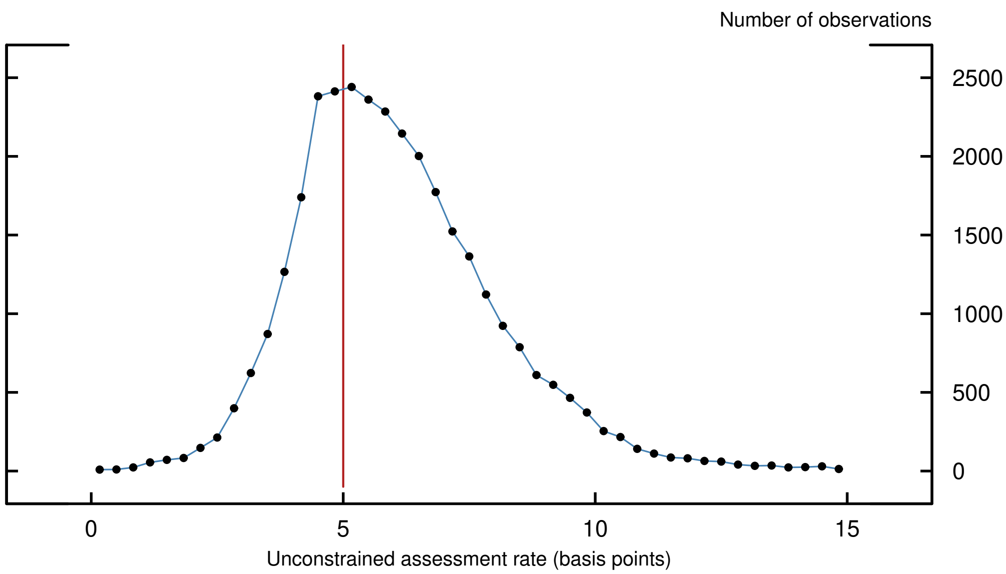


Figure 3.2: Distribution of Unconstrained Assessment Rates

Note: This figure shows the number of bank-quarter observations per bin of unconstrained assessment rates in our sample. Bins are 0.33 basis points wide and contain 716 observations on average. The vertical solid line identifies the minimum assessment rate of 5 basis points. The density of the running variable is continuous at the threshold of 5 basis points (Table 3.9).
Source: Consolidated Reports of Condition and Income (FFIEC 031 and 041) and Federal Reserve supervisory data.

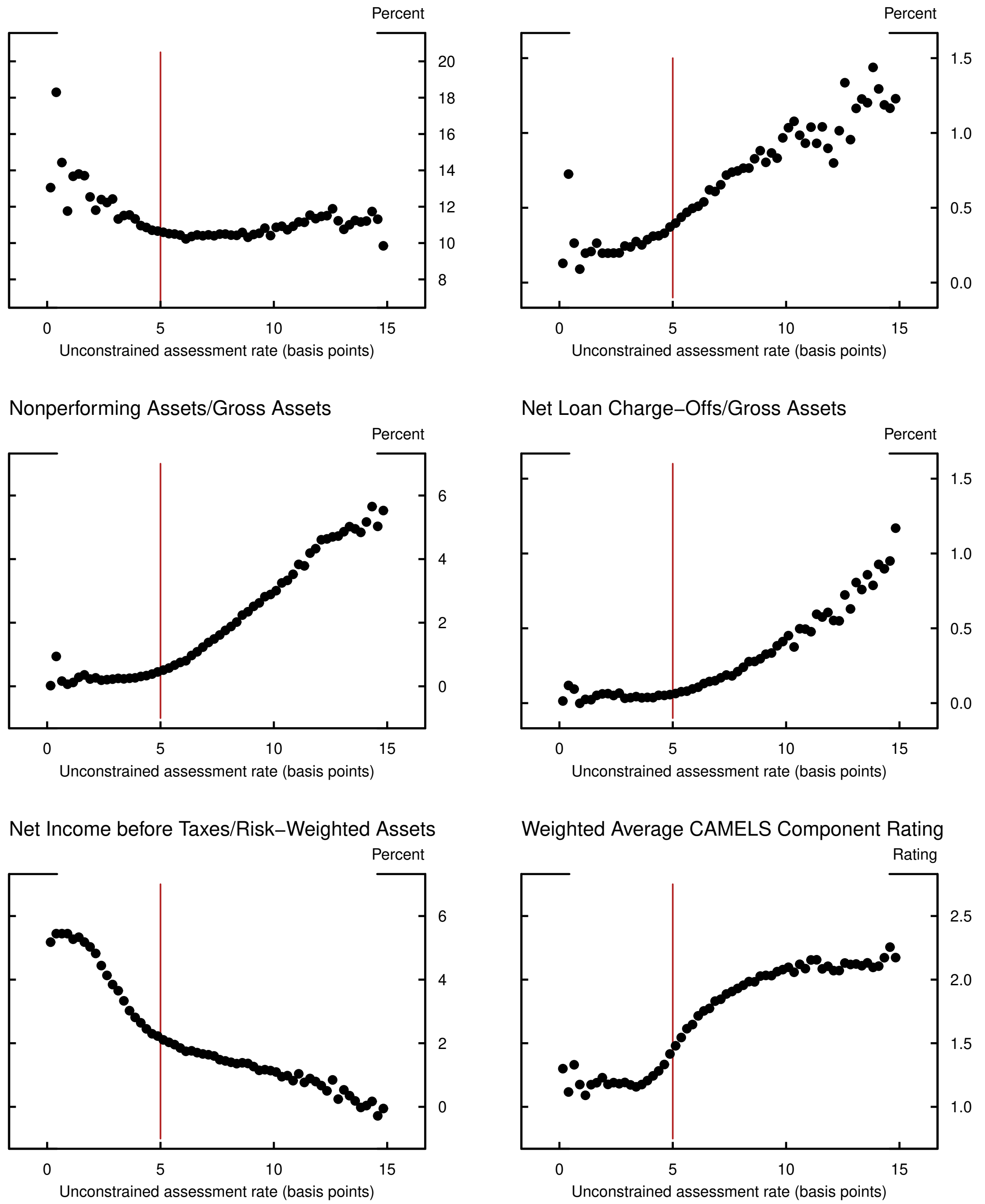


Figure 3.3: Smoothness Assumption on Assessment Rate Components

NOTE: This figure shows the mean values of covariates as a function of the running variable (unconstrained assessment rate). Mean values are measured in the same year-quarter as the running variable.
SOURCE: Consolidated Reports of Condition and Income (FFIEC 031 and 041) and Federal Reserve supervisory data.

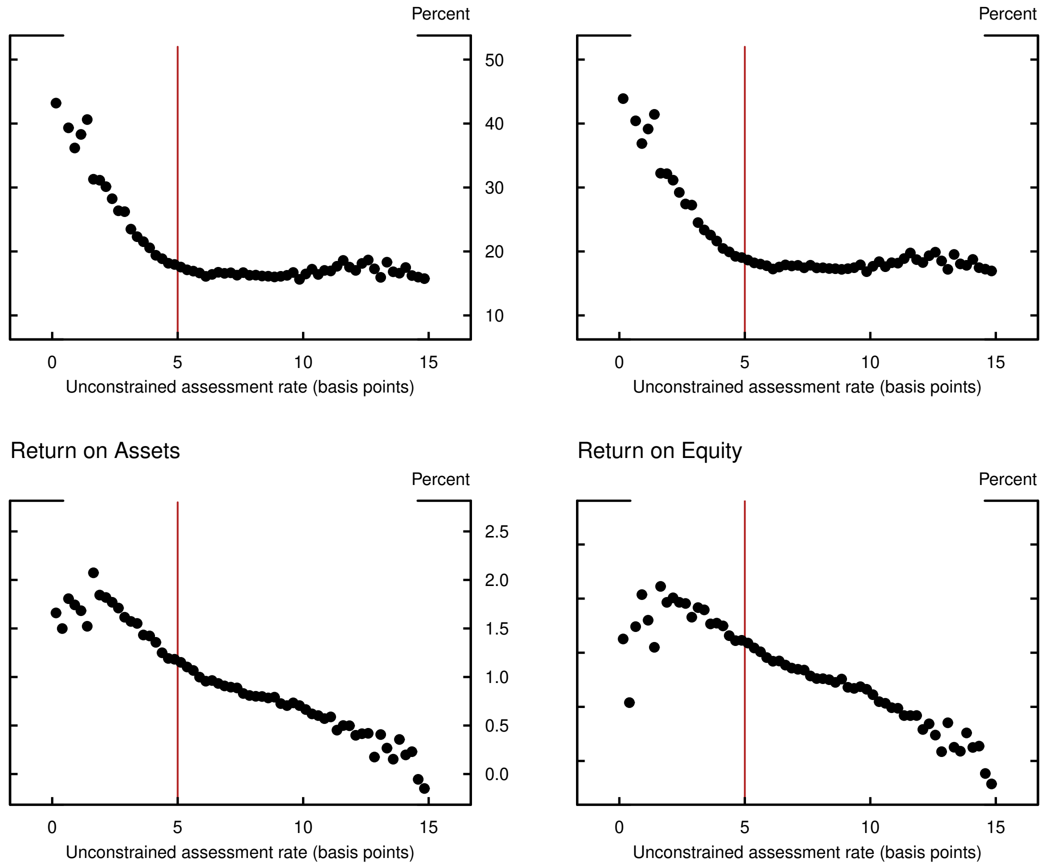


Figure 3.4: Smoothness Assumption on Covariates

Note: This figure shows the mean values of covariates as a function of the running variable (unconstrained assessment rate). Mean values are measured in the same year-quarter as the running variable, except ROA and ROE, which are measured in the previous year-quarter.
Source: Consolidated Reports of Condition and Income (FFIEC 031 and 041) and Federal Reserve supervisory data.

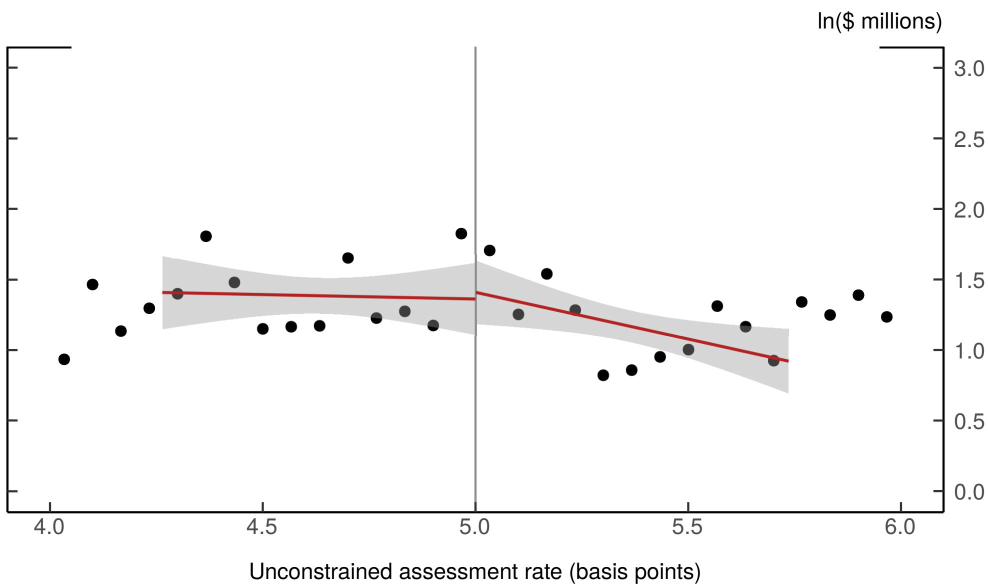


Figure 3.5: Assessment Rates and Excess Reserves

Note: This figure shows the relationship between unconstrained assessment rates (in the horizontal axis) and the natural logarithm of excess reserves (in the vertical axis). Unconstrained assessment rates are divided into thirty 0.67-basis point wide buckets. For each bucket, we plot the mean dependent variable and 95 percent confidence intervals. The vertical solid line identifies the minimum assessment rate of 5 basis points. The red lines on the left and on the right of the 5-basis point cutoff are predicted values from local linear regressions estimated with a bandwidth equal to 0.736, the optimal bandwidth from column 1 of Table 3.2.
Source: Consolidated Reports of Condition and Income (FFIEC 031 and 041) and Federal Reserve supervisory data.

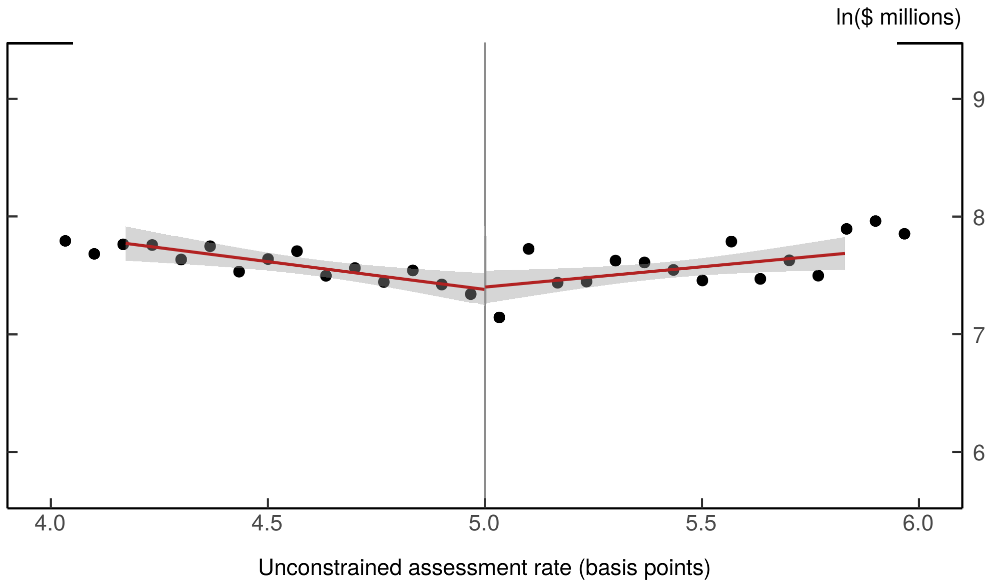


Figure 3.6: Assessment Rates and Federal Funds Sold

Note: This figure shows the relationship between unconstrained assessment rates (in the horizontal axis) and the natural logarithm of federal funds sold (in the vertical axis). Unconstrained assessment rates are divided into thirty 0.67-basis point wide buckets. For each bucket, we plot the mean dependent variable and 95 percent confidence intervals. The vertical solid line identifies the minimum assessment rate of 5 basis points. The red lines on the left and on the right of the 5-basis point cutoff are predicted values from local linear regressions estimated with a bandwidth equal to 0.829, the optimal bandwidth from column 1 of Table 3.3.
Source: Consolidated Reports of Condition and Income (FFIEC 031 and 041) and Federal Reserve supervisory data.

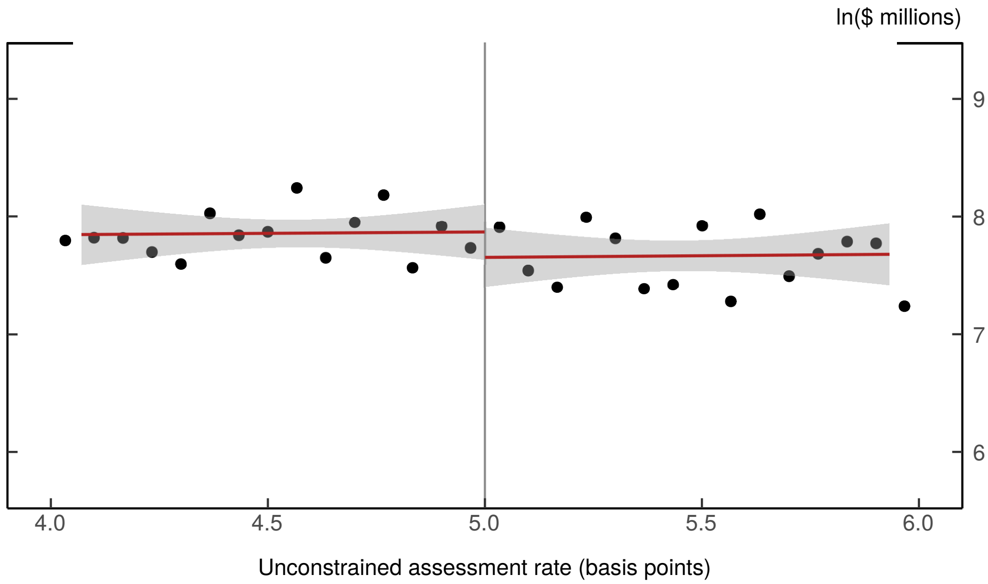


Figure 3.7: Assessment Rates and Federal Funds Purchased

Note: This figure shows the relationship between unconstrained assessment rates (in the horizontal axis) and the natural logarithm of federal funds purchased (in the vertical axis). Unconstrained assessment rates are divided into thirty 0.67-basis point wide buckets. For each bucket, we plot the mean dependent variable and 95 percent confidence intervals. The vertical solid line identifies the minimum assessment rate of 5 basis points. The lines on the left and on the right of the 5-basis point cutoff are predicted values from local linear regressions estimated with a bandwidth equal to 0.932, the optimal bandwidth from column 2 of Table 3.3.
Source: Consolidated Reports of Condition and Income (FFIEC 031 and 041) and Federal Reserve supervisory data.

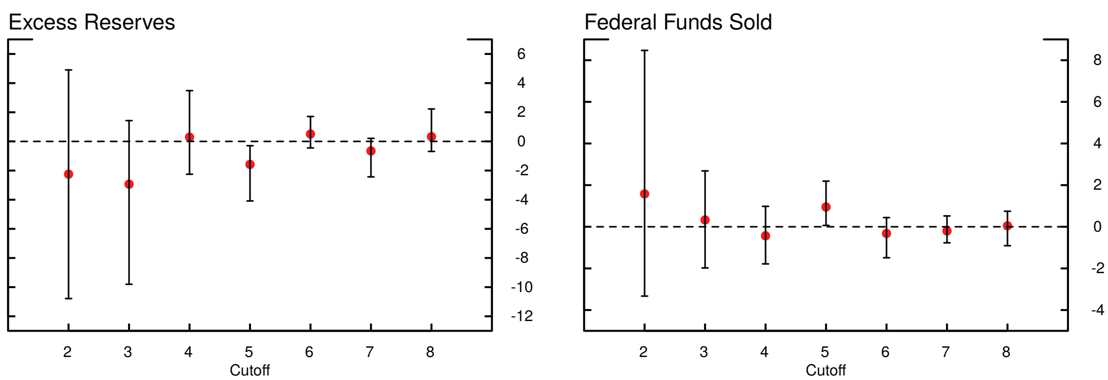


Figure 3.8: RKD Estimates with True and Placebo Cutoffs

NOTE: This figure shows RKD estimates using alternative cutoffs from 2 basis points to 8 basis points. The true cutoff is at 5 basis points. The left and right panels use the natural logarithms of excess reserves and federal funds sold as dependent variables, respectively. Red dots show our point estimates of treatment effects and the vertical lines show robust 95 percent confidence intervals. Table 3.11 in Appendix 3.D.2 shows the complete results.

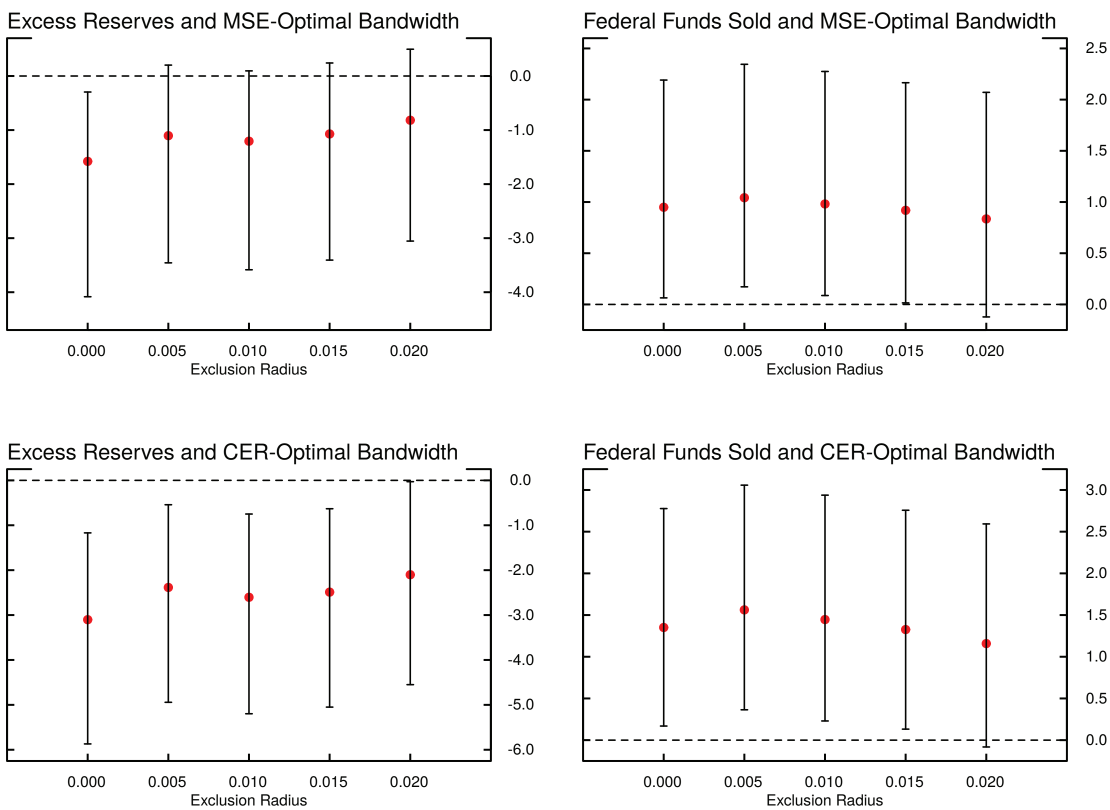


Figure 3.9: RKD Estimates Excluding Observations near the Cutoff

NOTE: This figure shows RKD estimates eliminating observations within a neighborhood of the cutoff ranging from 0 to 0.020 basis points to the left and to the right of the 5-basis point cutoff. The left and right panels use the natural logarithms of excess reserves and of federal funds sold as dependent variables. Red dots show our point estimates of treatment effects and the vertical lines show robust 95 percent confidence intervals. Table 3.12 in Appendix 3.D.3 shows the complete results.

	Obs.	Mean	Std. Dev.
Outcome variables			
Reserves ($ millions)	19,124	21.09	67.51
Excess reserves ($ millions)	19,124	5.46	37.30
Average reserves ($ millions)	19,124	21.58	64.73
Average excess reserves ($ millions)	18,995	5.42	35.55
Federal funds sold ($ millions)	32,384	3.56	16.20
Federal funds purchased ($ millions)	32,384	1.60	50.18
Repo ($ millions)	32,384	0.10	2.10
Reverse repo ($ millions)	32,384	4.72	22.72
Assignment variables			
Unconstrained assessment rate (b.p.)	32,384	6.06	2.08
Tier 1 leverage ratio (pct.)	32,384	10.77	3.30
Loans past due 30-89 days/Gross assets (pct.)	32,384	0.53	0.58
Nonperforming Assets/Gross assets (pct.)	32,384	1.04	1.05
Net loan charge-offs/Gross assets (pct.)	32,384	0.14	0.24
Net income before taxes/Risk-weighted assets (pct.)	32,384	2.04	1.14
Adjusted brokered deposit ratio (pct.)	32,384	0.00	0.00
Weighted average CAMELS component rating	32,384	1.62	0.40
Other bank characteristics			
Total capital ratio (pct.)	32,384	19.16	7.65
Tier 1 capital ratio (pct.)	32,384	18.04	7.67
Return on assets (pct.)	32,384	1.07	0.72
Return on equity (pct.)	32,384	9.86	6.66

Table 3.1: Summary Statistics

NOTE: Each observation is a bank-quarter pair and the sample period ranges from the second quarter of 2011 to the second quarter of 2016. The data are restricted to domestic commercial banks in FDIC's risk category 1 that have been open for more than five years, have total assets between $100 million and $5 billion, and are not subject to the unsecured debt adjustment, brokered deposit adjustment, and depository institution debt adjustment.
SOURCE: Consolidated Reports of Condition and Income (FFIEC 031 and 041) and Federal Reserve supervisory data.

	Local linear		Local quadratic	
	Quarter-end excess reserves (1)	Average excess reserves (2)	Quarter-end excess reserves (3)	Average excess reserves (4)
τ	-1.579	-1.694	-2.680	-2.814
Robust 95% CI	[-4.084, -0.296]	[-4.182, -0.452]	[-6.094, -0.301]	[-6.231, -0.488]
Robust p-value	0.023	0.015	0.030	0.022
N_-	3,131	3,082	4,523	4,456
N_+	3,300	3,244	5,607	5,483
h	0.736	0.727	1.307	1.281

Table 3.2: Effects of Assessment Rates on Bank Reserves

Note: Point estimators are constructed using local polynomial estimators with triangular kernel. Robust p-values are constructed using bias-correction with robust standard errors as derived in Calonico et al. (2014). N_- and N_+ are the number of observations effectively used above and below the 5-basis point cutoff out of 18,907 (columns 1 and 3) and 18,805 (columns 2 and 4) observations. h is the second generation data-driven MSE-optimal bandwidth selector proposed in Calonico et al. (2014).

Panel A: Federal funds				
	Local linear		Local quadratic	
	Federal funds sold (1)	Federal funds purchased (2)	Federal funds sold (3)	Federal funds purchased (4)
τ	0.949	-0.093	1.212	0.079
Robust 95% CI	[0.260, 1.638]	[-1.554, 1.652]	[0.170, 2.413]	[-2.261, 3.126]
Robust p-value	0.007	0.932	0.024	0.753
N_-	2,606	724	4,057	887
N_+	2,862	652	6,096	860
h	0.829	0.932	1.903	1.304

Panel B: Federal funds and securities repurchase agreements				
	Local linear		Local quadratic	
	Federal funds sold + reverse repo (1)	Federal funds purchased + repo (2)	Federal funds sold + reverse repo (3)	Federal funds purchased + repo (4)
τ	0.961	-0.023	1.196	-0.127
Robust 95% CI	[0.077, 2.201]	[-1.680, 1.718]	[0.164, 2.384]	[-3.194, 2.868]
Robust p-value	0.036	0.983	0.047	0.916
N_-	2,612	728	4,003	890
N_+	2,867	660	5,884	863
h	0.830	0.944	1.826	1.315

Table 3.3: Effects of Assessment Rates on Interbank Lending

NOTE: Point estimators are constructed using local polynomial estimators with triangular kernel. Robust p-values are constructed using bias-correction with robust standard errors as derived in Calonico et al. (2014). N_- and N_+ are the number of observations effectively used above and below the 5-basis point cutoff out of 15,272 (columns 1 and 3) and 3,098 (columns 2 and 4) observations. h is the second generation data-driven MSE-optimal bandwidth selector proposed in Calonico et al. (2014).

Bandwidth selection procedure	RKD treatment effect	Robust 95% CI	Robust p-value	N_{-}	N_{+}	h_{-}	h_{+}
Panel A: Excess reserves as dependent variable							
Common CER-optimal	-3.103	[-5.869, -1.170]	0.003	2,198	2,284	0.500	0.500
Common MSE-optimal	-1.579	[-4.084, -0.296]	0.023	3,131	3,300	0.736	0.736
Different CER-optimal	-2.559	[-4.875, -0.931]	0.004	2,551	2,754	0.583	0.611
Different MSE-optimal	-0.980	[-3.143, 0.181]	0.081	3,536	3,979	0.858	0.899
Panel B: Federal funds sold as dependent variable							
Common CER-optimal	1.351	[0.169, 2.779]	0.027	1,892	1,951	0.568	0.568
Common MSE-optimal	0.949	[0.064, 2.191]	0.038	2,606	2,862	0.829	0.829
Different CER-optimal	1.271	[0.115, 2.625]	0.032	1,797	2,339	0.540	0.676
Different MSE-optimal	0.966	[0.085, 2.138]	0.034	2,515	3,383	0.788	0.986

Table 3.4: Effects of Assessment Rates with Alternative Bandwidths

NOTE: Point estimators are constructed using local polynomial estimators with triangular kernel. Robust p-values are constructed using bias-correction with robust standard errors as derived in Calonico et al. (2014). Panels A and B use, respectively, the natural logarithms of excess reserves and federal funds sold measured in millions of dollars as the dependent variable. N_{-} and N_{+} are the number of observations effectively used above and below the 5-basis point cutoff out of 18,907 (Panel A) and 15,272 (Panel B) observations. h_{-} and h_{+} are the second generation data-driven MSE-optimal bandwidth selectors proposed in Calonico et al. (2014) above and below the 5-basis point cutoff.

3.B IOER Rate and Federal Funds Rates

In this appendix, we discuss the relationship between the IOER rate and the rates on overnight federal funds transactions. In principle, the IOER rate should serve as a floor for interbank loans. However, throughout our sample period, the IOER rate stayed above the EFFR, which is a volume-weighted median of the rate on overnight federal funds transactions. The EFFR is often below the IOER rate primarily due to loan supply from other financial institutions that cannot hold reserves at the Federal Reserve. These institutions, which include government-sponsored enterprises (for example, Fannie Mae, Freddie Mac, and Federal Home Loan Banks), collectively account for about three- quarters of total interbank lending in the federal funds market since the Global Financial Crisis. Banks that do not have reserve accounts and can only earn IOER at a discount also contribute to the downward pressure on the EFFR, but most likely have a weaker effect than GSEs and other financial institutions because of their smaller size. See Bech and Klee (2011) for a discussion on the topic.

Even though the IOER may exceed the EFFR, banks may find it profitable to supply interbank loans for various reasons. First, many small banks without reserve accounts at a Federal Reserve Bank can only receive IOER via accounts managed by correspondent banks that charge a fee for this service. Accounting for those fees, holding excess reserves becomes less profitable as these banks cannot earn the full IOER rate. Second, these banks can more easily lend to correspondent banks without transferring funds (Afonso et al., 2011). Thus, small banks with no reserve accounts—about two-fifths of the banks in our sample—may find interbank loans to correspondent banks less costly or more convenient than holding excess reserves.

Third, in the case of banks with reserve accounts, we argue that they likely extend interbank loans at rates above the IOER rate. The EFFR is calculated based on transactions that include banks that can earn IOER and other financial institutions that cannot earn IOER. Because the EFFR is pressured downwards by institutions that cannot earn IOER, banks that supply interbank loans in lieu of earning IOER likely do so at a rate higher than the IOER rate. Indeed, interbank rates can be above the IOER rate due to search costs incurred by borrowing entities or increased demand for funds late in the trading day (Afonso and Lagos, 2015; Kim et al., 2020).

3.C Assessment Rates

This appendix provides more details on the assessment rates discussed in Section 3.1 and explains how we calculate them for this paper.

3.C.1 Initial Base Assessment Rate

Table 3.5 describes how the capital ratios and the CAMELS composite rating of a bank determine its risk category. Risk categories range from category 1 to 4, with risk category 1 generally containing well-capitalized banks with good ratings (CAMELS of 1 or 2) and risk category 4 generally containing undercapitalized banks with bad ratings (CAMELS of 4 or 5).

Based on the risk category of the bank, the FDIC assigns it an initial base assessment rate. Table 3.6 shows the rates charged during our sample period, from April 1, 2011, to June 30, 2016. The FDIC assigns to each risk category 1 bank an initial base assessment rate that ranges from 5 to 9 basis points during this period. Risk category 2, 3, and 4 banks are assessed initial base assessment rates of 14, 23, and 35 basis points, respectively, regardless of their characteristics.

The FDIC computes the rate of risk category 1 banks by calculating the sum of risk measures at the bank level multiplied by coefficients derived from an econometric model of bank failures (Federal Deposit Insurance Corporation, 2011). These measures and their coefficients are outlined in Table 3.7. The weighted average CAMELS component rating is calculated by taking the weighted sum of each of the component ratings, using the weights outlined in Table 3.8. The sum of the risk measures multiplied by the coefficients from Table 3.7 is also added to a uniform amount, which is equal to 4.861 basis points for our sample period. We define the total as the unconstrained initial base assessment rate, and simply refer to it as the unconstrained assessment rate.

The unconstrained assessment rate of a risk category 1 bank is constrained by the minimum and maximum rates shown in the first column of Table 3.6. The constrained assessment rate is equal to the minimum rate of 5 basis points if the unconstrained assessment rate is below this minimum and it is equal to the maximum rate of 9 basis points if the unconstrained assessment rate is above this maximum. As shown by the solid line in Figure 3.1, this rule creates a relationship between the constrained and the unconstrained assessment rates that is flat to the left of 5 basis points, increasing with a slope equal to 1 between 5 and 9 basis points, and also flat to the right of 9 basis points.

3.C.2 Adjustments to the Unconstrained Assessment Rate

After a bank's unconstrained assessment rate is calculated, this rate may be adjusted downward for unsecured debt (UDA) and upward for brokered deposits (BDA) and for debt issued by other institutions (DIDA). The UDA of a bank is calculated by adding 40 basis points to the unconstrained assessment rate and multiplying this sum by the ratio of the bank's long-term unsecured debt to its assessment base. This amount, limited to a maximum equal to the lesser of 5 basis points and 50 percent of the bank's unconstrained assessment rate, is subtracted from the unconstrained assessment rate. Conversely, the BDA only applies to banks in risk categories 2 to 4 and whose ratio of brokered deposits to domestic deposits is greater than 10 percent.

This adjustment is calculated as 25 basis points times the ratio of the difference between brokered deposits and 10 percent of its domestic deposits to its assessment base. This amount, limited to a minimum of zero and a maximum of 10 basis points, is added to the unconstrained assessment rate. Lastly, the DIDA is a 50 basis point charge on the amount of long-term unsecured debt that was issued by another insured depository institution and that exceeds 3 percent of the bank's tier 1 capital. The rate that results from these three adjustments, and which is actually charged to banks, is defined as the total base assessment rate. We also refer to this rate as the assessment rate throughout the paper.

Among these three adjustments, the UDA is the only one that affects our estimates of the effects of deposit insurance premiums on bank behavior. In this paper, we use the change in the slope of the total base assessment rate of risk category 1 banks as a function of the unconstrained assessment rate to identify these effects. Thus, the BDA does not affect these estimates because this adjustment only applies to banks in risk categories 2 to 4. Moreover, the DIDA does not affect the estimates because it does not depend on the unconstrained assessment rate.

The UDA attenuates the changes in slope of the constrained assessment rate as a function of the unconstrained assessment rate, thereby affecting the economic interpretation of our coefficient estimates. The UDA attenuates the changes at 5 basis points, shown in Figure 3.10, because it is in absolute terms an increasing function of the constrained assessment rate. Indeed, the change in slope at 5 basis points is largest when the UDA is equal to zero (solid line) and is smallest when the UDA reaches its cap of 50 percent of the unconstrained assessment rate (dashed line). For this reason, the economic effect implied by coefficient estimates under the assumption that the UDA is equal to zero (i.e. the slope of the total base assessment rate as a function of the unconstrained assessment rate changes from 1 to zero) is a lower bound for the effect implied by those estimates without this assumption. Similarly, the economic effect implied by coefficient estimates under the assumption that the UDA is the highest possible (i.e. the slope of the total base assessment rate as a function of the unconstrained assessment rate changes from 0.5 to zero) is a higher bound for the effect implied by those estimates without this assumption and is twice as large as the lower bound.

3.C.3 Calculation of Assessment Rates

The rule that determines the method for calculating assessment rates is described in Federal Deposit Insurance Corporation (2011). The FDIC also publishes on their website a calculator that illustrates how a bank's assessment rate is determined. The calculator, maintained in the form of a spreadsheet, is designed to help banks understand how their assessment rates are calculated and to help banks simulate the impact of changes in their characteristics on their rates. Its downside, however, is that it is designed to show rates for a single bank at a time. To overcome this challenge, we first use the FDIC's documentation and our own data set to compute the unconstrained assessment rate for all small banks during our sample period. We then compare our data-driven rate to the calculator's rate for a random selection of banks and verify that the numbers are exactly the same.

We also compare our assessment rate calculations to Call Report data on dollar amounts of FDIC fees. Specifically, we analyze whether the reported assessment fees divided by the assessment base is consistent with our calculated assessment rates. The results are similar on balance but not always the same, indicating that actual FDIC charges may differ from the amounts that banks report in the Call Reports for various reasons. For instance, payments might be delayed or banks and the FDIC may disagree on the amount charged.

3.D Additional Validation and Falsification Results

3.D.1 Evidence on the Smoothness Assumption

In this appendix, we present additional evidence that the assumptions of the RKD are satisfied in our setting. Table 3.9 presents tests of the null hypothesis that the density of the running variable is continuous at the cutoff of 5 basis points. The three columns show the results of tests proposed by Cattaneo et al. (2018, 2020) under different specifications. The p-values are large in all three columns, indicating that the null hypothesis is never rejected by these tests. Thus, these results support the validity of the RKD in our setting.

We now test formally whether the conditional expectations of the ten covariates shown in Figures 3.3 and 3.4 are twice continuously differentiable around the threshold of 5 basis points. We estimate treatment effects on those covariates using the estimator $\hat{\tau}(h_{IT})$ and the cutoff of 5 basis points. Table 3.10 shows the results of these tests using our preferred specification for the estimator (local-linear). We find that the estimate of the robust 95-percent confidence interval for all covariates includes zero and the robust p-value does not allow us to reject the null hypothesis that $\tau = 0$. This finding supports the smoothness assumption of our RKD.

3.D.2 Placebo Cutoffs

We next present estimates of the effects of assessment rates on excess reserve balances and amounts of federal funds sold using alternative cutoff points. Table 3.11 shows the complete results summarized in Figure 3.8.

3.D.3 Sensitivity to Observations Near the Cutoff

We examine whether the estimates in Tables 3.2 and 3.3 change materially if we drop observations close to the 5-basis point cutoff. Table 3.12 shows the complete results summarized in Figure 3.9.

3.E Additional Results

In this appendix, we provide evidence that banks' unconstrained assessment rates are correlated over time. This evidence supports an assumption that we introduce in Section 3.4, namely that banks consider the rates in t reliable approximations of their rates in $t+1$. To evaluate this assumption, we use ordinary least squares (OLS) to estimate the following equation:

$$UAR_{i,t+1} = \alpha \times UAR_{i,t} + \nu_i + \varphi_t + \varepsilon_{it}, \tag{3.6}$$

where $UAR_{i,t}$ is the unconstrained assessment rate of bank i in period t measured in basis points, ν_i and φ_t are bank and time fixed effects, and ε_{it} an idiosyncratic shock. Standard errors are clustered at the bank level. α is the coefficient of interest, and we test the null hypothesis that $\alpha = 0$. A rejection of the null hypothesis with an estimate of α close to 1 indicates that unconstrained assessment rates are correlated over time within banks, and we interpret these findings as evidence in favor of our assumption.

Table 3.13 presents estimates of equation (3.6). In column 1, we use the same sample from Section 3.4. In columns 2 to 4, we restrict the sample to bank-quarter observations such that $UAR_{i,t}$ belongs to the interval $(5-h, 5+h)$, where h ranges between 2 basis points (column 2) and 0.5 basis points (column 4). We estimate equation (3.6) with different bandwidths to examine whether evidence that assessment rates are correlated over time depends on the distance between rates and the 5-basis point cutoff.

The estimates of α in the four columns are positive and statistically significant, rejecting the hypothesis of no correlation over time in assessment rates and, therefore, supporting our assumption. Although the estimate decreases as we narrow the bandwidth (going from column 1 to 4), the results show that rates remain correlated over time within banks even when we employ a bandwidth of $h = 0.5$—the narrowest used in Section 3.4. These results support our assumption that banks consider the rates in t to be reliable approximations of their rates in $t + 1$.

3.F Appendix Figures and Tables

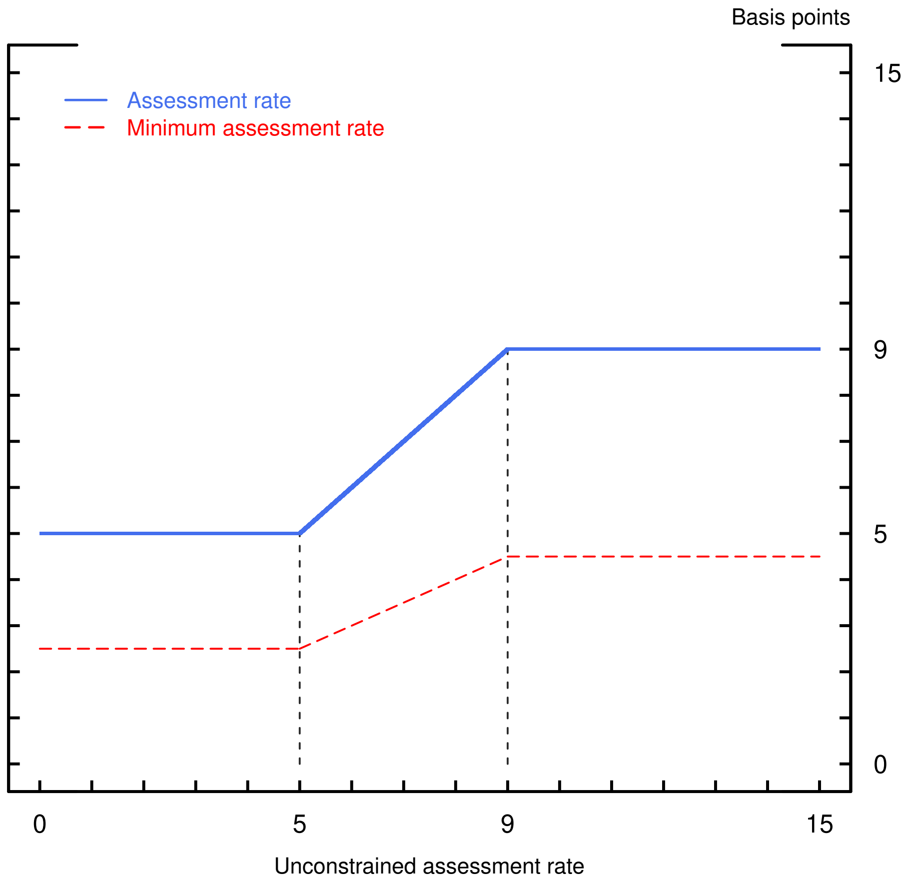


Figure 3.10: Kinks at Assessment Rate with Unsecured Debt Adjustment

NOTE: The solid line shows the constrained assessment rate as a function of the unconstrained assessment rate for insured risk category 1 banks between April 1, 2011, and June 30, 2016, with total assets below $10 billion. Newly insured institutions (those that became insured within five years) are subject to different rates and are not included in the analysis. Assessment rates are measured in basis points.
SOURCE: Federal Deposit Insurance Corporation (2011).

	Supervisory group**		
Capital group*	A	B	C
1 (Well capitalized)	I	II	III
2 (Adequately capitalized)	II	II	III
3 (Undercapitalized)	III	III	IV

Table 3.5: Risk Category Schedule

NOTE: * Well capitalized banks are defined as banks with total risk-based capital ratio equal to or greater than 10 percent, tier 1 risk-based capital ratio equal to or greater than 6 percent, and tier 1 leverage capital ratio equal to or greater than 5 percent; adequately capitalized banks are defined as banks that are not well capitalized and have total risk-based capital ratio equal to or greater than 8 percent, tier 1 risk-based capital ratio equal to or greater than 4 percent, and tier 1 leverage capital ratio equal to or greater than 4 percent; and undercapitalized banks are defined as banks that are neither well capitalized nor adequately capitalized.
** Supervisory group A generally includes banks with CAMELS composite ratings of 1 or 2, supervisory group B generally includes banks with a CAMELS composite rating of 3, and supervisory group C generally includes banks with CAMELS composite ratings of 4 or 5.
SOURCE: Federal Deposit Insurance Corporation (2011).

Table 3.6: Initial Base Assessment Rate Schedule

Risk category	I	II	III	IV
Initial base assessment rate	5 to 9	14	23	35
Unsecured debt adjustment	-4.5 to 0	-5 to 0	-5 to 0	-5 to 0
Brokered deposit adjustment	N/A	0 to 10	0 to 10	0 to 10
Total base assessment rate	2.5 to 9	9 to 24	18 to 33	30 to 45

NOTE: All amounts for all categories are in basis points annually. Total base assessment rates do not include the depository institution debt adjustment.
SOURCE: Federal Deposit Insurance Corporation (2011).

Risk measures	Coefficients
Tier 1 leverage ratio	-0.056
Loans past due 30-89 days / gross assets	0.575
Nonperforming assets / gross assets	1.074
Net loan charge-offs / gross assets	1.210
Net income before taxes / risk-weighted assets	-0.764
Adjusted brokered deposit ratio	0.065
Weighted average CAMELS component rating	1.095

Table 3.7: Risk Measures and Coefficients

NOTE: Ratios are expressed as percentages and pricing multipliers are rounded to three decimal places.
SOURCE: Federal Deposit Insurance Corporation (2011).

Component	Weight (percent)
Capital adequacy	25
Asset quality	20
Management administration	25
Earnings	10
Liquidity	10
Sensitivity to market risk	10

Table 3.8: Weighted Average CAMELS Component Rating

Note: Each numerical rating is a round number between 1 and 5. The weighted average component rating is computed by multiplying the rating by the weight, and summing across the six categories. The results are rounded to three decimal places for unconstrained assessment rate calculation.
Source: Federal Deposit Insurance Corporation (2011).

	Unrestricted inference with distinct bandwidths (1)	Unrestricted inference with identical bandwidths (2)	Restricted inference with identical bandwidths (3)
h_-	0.823	1.427	0.617
h_+	0.978	1.427	0.617
N_-	5,712	8,062	4,477
N_+	6,940	9,790	4,448
p-value	0.511	0.587	0.907

Table 3.9: Density Tests of Assessment Fees

Note: This table shows tests of the null hypothesis that the density of the running variable is continuous at the cutoff of 5 basis points. h_- and h_+ denote the estimator bandwidth on the left and on the right of the cutoff, respectively. N_- and N_+ denote the effective number of observations used above and below the 5-basis point cutoff out of 32,384 observations. Density test p-values are computed using Gaussian distributional approximation to bias-corrected local-linear polynomial estimator with triangular kernel and robust standard errors. Column 1 shows results of unrestricted inference with two distinct bandwidths, column 2 shows results of unrestricted inference with one common bandwidth, and column 3 shows results of restricted inference with one common bandwidth. See Cattaneo et al. (2018, 2020) for methodological and implementation details.

	Tier 1 leverage ratio	Loans past due to gross assets ratio	Nonperf. assets to gross assets ratio	Net loan chg-offs to gross assets ratio
	(1)	(2)	(3)	(4)
RKD treat. eff.	-0.136	0.032	-0.049	0.015
Robust 95% CI	[-1.788, 0.952]	[-0.098, 0.120]	[-0.334, 0.046]	[-0.037, 0.039]
Robust p-value	0.550	0.838	0.137	0.950
N_{-}	4,569	6,883	4,209	6,087
N_{+}	4,553	7,655	4,159	6,404
h	0.630	1.083	0.575	0.902

	NIBT to R-W assets ratio	Weighted average CAMELS	Total capital ratio	Tier 1 capital ratio
	(5)	(6)	(7)	(8)
RKD treat. eff.	-0.067	-0.034	-0.589	-0.526
Robust 95% CI	[-0.583, 0.139]	[-0.229, 0.082]	[-4.806, 1.245]	[-4.720, 1.332]
Robust p-value	0.229	0.354	0.249	0.272
N_{-}	4,053	4,805	4,320	4,338
N_{+}	4,001	4,824	4,287	4,306
h	0.554	0.669	0.592	0.595

	Return on assets	Return on equity
	(9)	(10)
RKD treat. eff.	-0.155	-1.498
Robust 95% CI	[-0.511, 0.006]	[-4.804, 0.360]
Robust p-value	0.056	0.092
N_{-}	4,665	4,891
N_{+}	4,643	4,950
h	0.646	0.687

Table 3.10: Treatment Effects on Covariates

NOTE: This table shows estimates of treatment effects on covariates using a cutoff of 5 basis points. Point estimators are constructed using local-quadratic polynomial estimators with triangular kernel. Robust p-values are constructed using bias-correction with robust standard errors as derived in Calonico et al. (2014). h is the second generation data-driven MSE-optimal bandwidth selector from Calonico et al. (2014). N_{-} and N_{+} denote the effective number of observations on the left and on the right of the cutoff, respectively. All variables are measured in the same year-quarter as the running variable, except ROA and ROE, which are measured in the previous year-quarter.

Alternative cutoff (b.p.)	RKD treatment effect	Robust 95% CI	Robust p-value	N_-	N_+	h
		Panel A: Excess reserves as dependent variable				
2	-2.247	[-10.784, 4.904]	0.463	87	337	0.929
3	-2.934	[-9.802, 1.428]	0.144	359	1,046	0.754
4	0.292	[-2.250, 3.484]	0.673	1,361	3,073	0.806
5	-1.579	[-4.084, -0.296]	0.023	3,131	3,300	0.736
6	0.506	[-0.448, 1.710]	0.252	5,244	4,048	1.197
7	-0.652	[-2.433, 0.209]	0.099	3,271	2,199	0.942
8	0.323	[-0.687, 2.227]	0.301	2,314	1,297	1.002
		Panel B: Federal funds sold as dependent variable				
2	1.579	[-3.330, 8.471]	0.393	68	210	0.853
3	0.327	[-1.978, 2.683]	0.767	302	1,206	1.028
4	-0.434	[-1.783, 0.981]	0.570	1,044	2,388	0.827
5	0.949	[0.064, 2.191]	0.038	2,606	2,862	0.829
6	-0.322	[-1.487, 0.442]	0.288	2,807	2,426	0.812
7	-0.199	[-0.770, 0.521]	0.706	3,442	2,200	1.154
8	0.046	[-0.909, 0.746]	0.846	2,354	1,299	1.148

Table 3.11: Effects of Assessment Rates Using Alternative Cutoffs

NOTE: This table shows estimates of treatment effects on covariates using alternative cutoff points. Panel A uses the natural logarithm of excess reserves measured in millions of dollars as the dependent variable, and Panel B uses the natural logarithm of federal funds sold measured in millions of dollars. Point estimators are constructed using local-quadratic polynomial estimators with triangular kernel. Robust p-values are constructed using bias-correction with robust standard errors as derived in Calonico et al. (2014). h is the second generation data-driven MSE-optimal bandwidth selector from Calonico et al. (2014). N_- and N_+ denote the effective number of observations on the left and on the right of the cutoff, respectively.

Exclusion radius (b.p.)	RKD treatment effect	Robust 95% CI	Robust p-value	N_-	N_+	h	Observ. excluded on left	Observ. excluded on right
Panel A: Excess reserves as dependent variable and MSE-optimal bandwidth								
0.000	-1.579	[-4.084, -0.296]	0.023	3,313	3,300	0.736	0	0
0.005	-1.104	[-3.457, 0.202]	0.081	3,216	3,406	0.766	21	29
0.010	-1.206	[-3.585, 0.095]	0.063	3,178	3,367	0.761	38	45
0.015	-1.072	[-3.406, 0.241]	0.089	3,222	3,418	0.777	60	64
0.020	-0.818	[-3.054, 0.496]	0.158	3,278	3,489	0.800	78	84
Panel B: Federal funds sold as dependent variable and MSE-optimal bandwidth								
0.000	0.949	[0.064, 2.191]	0.038	2,606	2,862	0.829	0	0
0.005	1.042	[0.172, 2.345]	0.023	2,552	2,782	0.810	13	22
0.010	0.981	[0.087, 2.275]	0.034	2,570	2,808	0.824	27	38
0.015	0.919	[0.015, 2.165]	0.047	2,550	2,793	0.823	44	50
0.020	0.835	[-0.122, 2.071]	0.081	2,551	2,802	0.832	63	68
Panel C: Excess reserves as dependent variable and CER-optimal bandwidth								
0.000	-3.103	[-5.870, -1.170]	0.003	2,198	2,284	0.500	0	0
0.005	-2.385	[-4.943, -0.543]	0.015	2,256	2,337	0.520	21	29
0.010	-2.604	[-5.198, -0.750]	0.009	2,227	2,311	0.517	38	45
0.015	-2.489	[-5.050, -0.632]	0.012	2,252	2,339	0.528	60	64
0.020	-2.102	[-4.550, -.0292]	0.026	2,302	2,384	0.544	78	84
Panel D: Federal funds sold as dependent variable and CER-optimal bandwidth								
0.000	1.351	[0.169, 2.777]	0.027	1,892	1,951	0.568	0	0
0.005	1.562	[0.364, 3.058]	0.013	1,838	1,882	0.556	13	22
0.010	1.446	[0.230, 2.938]	0.022	1,853	1,896	0.565	27	38
0.015	1.326	[0.132, 2.757]	0.031	1,829	1,880	0.564	44	50
0.020	1.158	[-0.082, 2.593]	0.066	1,835	1,889	0.570	63	68

Table 3.12: Effects of Assessment Rates Excluding Observations Near the Cutoff

NOTE: This table shows estimates of treatment effects on covariates when dropping observations near the 5 basis point cutoff. Panels A and C use the natural logarithm of excess reserves measured in millions of dollars as the dependent variable, and Panel B and D use the natural logarithm of federal funds sold measured in millions of dollars. Panels A and B use MSE-optimal bandwidths, and Panels C and D use CER-optimal bandwidths. Point estimators are constructed using local-quadratic polynomial estimators with triangular kernel. Robust p-values are constructed using bias-correction with robust standard errors as derived in Calonico et al. (2014). h is the second generation data-driven MSE-optimal bandwidth selector from Calonico et al. (2014). N_- and N_+ denote the effective number of observations used above and below the 5-basis point cutoff out of 18,907 (Panels A and C) and 15,272 (Panels B and D) observations.

	$h = 20$	$h = 2$	$h = 1$	$h = 0.5$
	(1)	(2)	(3)	(4)
$UAR_{i,t}$	0.800**	0.722**	0.705**	0.611**
	(0.018)	(0.017)	(0.028)	(0.054)
Observations	32,075	22,080	13,466	7,233
Banks	2,410	1,996	1,553	1,180
R-squared	0.78	0.68	0.65	0.61

Table 3.13: Correlation of Assessment Rates over Time

NOTE: This table shows OLS estimates of equation (3.6). The dependent variable is the unconstrained assessment rate of bank i in quarter $t + 1$ measured in basis points. h is the bandwidth that defines which observations on the left and on the right of the 5-basis point cutoff are used in the regression. All columns include bank and time fixed effects.

3.G Bibliography

Acharya, Sankarshan and Jean-Francois Dreyfus (1989) "Optimal Bank Reorganization Policies and the Pricing of Federal Deposit Insurance," *Journal of Finance*, 44 (5), 1313–1333.

Acharya, Viral V., Joao A. C. Santos, and Tanju Yorulmazer (2010) "Systemic Risk and Deposit Insurance Premiums," *Federal Reserve Bank of New York Economic Policy Review*, 16 (1), 89–99.

Afonso, Gara, Roc Armenter, and Benjamin Lester (2019) "A Model of the Federal Funds Market: Yesterday, Today, and Tomorrow," *Review of Economic Dynamics*, 33, 177–204.

Afonso, Gara, Kyungmin Kim, Antoine Martin, Ed Nosal, Simon Potter, and Sam Schulhofer-Wohl (2020) "Monetary Policy Implementation with an Ample Supply of Reserves," Federal Reserve Bank of Atlanta Working Paper Series, No. 2020-2.

Afonso, Gara, Anna Kovner, and Antoinette Schoar (2011) "Not Frozen: The Federal Funds Market in the Financial Crisis," *Journal of Finance*, 66 (4), 1109–1139.

Afonso, Gara and Ricardo Lagos (2015) "Trade Dynamics in the Market for Federal Funds," *Econometrica*, 83 (1), 263–313.

Allen, Franklin, Elena Carletti, and Robert Marquez (2015) "Deposits and Bank Capital Structure," *Journal of Financial Economics*, 118 (3), 601–619.

Armenter, Roc and Benjamin Lester (2017) "Excess Reserves and Monetary Policy Implementation," *Review of Economic Dynamics*, 23, 212–235.

Banegas, Ayelen and Manjola Tase (2020) "Reserve Balances, the Federal Funds Market and Arbitrage in the New Regulatory Framework," *Journal of Banking and Finance*, 118, 105893.

Basten, Christoph and Mike Mariathasan (2020) "Interest Rate Pass-Through and Bank Risk-Taking under Negative-Rate Policies with Tiered Remuneration of Central Bank Reserves," Swiss Finance Institute Research Paper No. 20-98.

Bech, Morten L. and Elizabeth Klee (2011) "The Mechanics of a Graceful Exit: Interest on Reserves and Segmentation in the Federal Funds Market," *Journal of Monetary Economics*, 58 (5), 415–431.

Boyd, John H., Chun Chang, and Bruce D. Smith (2002) "Deposit Insurance: A Reconsideration," *Journal of Monetary Economics*, 49 (6), 1235–1260.

Buser, Stephen A., Andrew H. Chen, and Edward J. Kane (1981) "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital," *Journal of Finance*, 36 (1), 51–60.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik (2017) "Rdrobust: Software for Regression Discontinuity Designs," *Stata Journal*, 17 (2), 372–404.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik (2014) "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82 (6), 2295–2326.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber (2015) "Inference on Causal Effects in a Generalized Regression Kink Design," *Econometrica*, 83 (6), 2453–2483.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma (2018) "Manipulation Testing Based on Density Discontinuity," *Stata Journal*, 18 (1), 234–261.

——— (2020) "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 115 (531), 1449–1455.

Cattaneo, Matias, Nicolas Idrobo, and Rocio Titiunik (2018) *A Practical Introduction to Regression Discontinuity Designs: Volume I*, forthcoming, *Cambridge Elements:*

Quantitative and Computational Methods for Social Science, Cambridge University Press.

Chan, Yuk-Shee, Stuart I. Greenbaum, and Anjan V. Thakor (1992) “Is Fairly Priced Deposit Insurance Possible?” *Journal of Finance*, 47 (1), 227–245.

Copeland, Adam, Darrell Duffie, and Yilin Yang (2021) “Reserves Were Not So Ample After All,” working paper.

Correa, Ricardo, Wenxin Du, and Gordon Liao (2020) “U.S. Banks and Global Liquidity,” Becker Friedman Institute working paper No. 2020-89.

Craine, Roger (1995) “Fairly Priced Deposit Insurance and Bank Charter Policy,” *Journal of Finance*, 50 (5), 1735–1746.

d’Avernas, Adrien and Quentin Vandeweyer (2021) “Intraday Liquidity and Money Market Dislocations,” working paper.

Dávila, Eduardo and Itay Goldstein (2020) “Optimal Deposit Insurance,” working paper.

Diamond, Douglas W. and Philip H. Dybvig (1983) “Bank Runs, Deposit Insurance, and Liquidity,” *Journal of Political Economy*, 91 (3), 401–419.

Duffie, Darrell, Robert Jarrow, Amiyatosh Purnanandam, and Wei Yang (2003) “Market Pricing of Deposit Insurance,” *Journal of Financial Services Research*, 24 (2-3), 93–119.

Duffie, Darrell and Arvind Krishnamurthy (2016) “Passthrough Efficiency in the Fed’s New Monetary Policy Setting,” in Babson, Richard A. ed. *Designing Resilient Monetary Policy Frameworks for the Future, A Symposium Sponsored by the Federal Reserve Bank of Kansas City*, 21–102, Jackson Hole, Wyoming: Federal Reserve Bank of Kansas City.

Duquerroy, Anne, Adrien Matray, and Farzad Saidi (2020) "Sticky Deposit Rates and Allocative Effects of Monetary Policy," working paper.

Federal Deposit Insurance Corporation (2011) "Assessments, Large Bank Pricing," Federal Register, 76(38), Friday, February 25.

Giammarino, Ronald M., Tracy R. Lewis, and David E. M. Sappington (1993) "An Incentive Approach to Banking Regulation," *Journal of Finance*, 48 (4), 1523–1542.

Heider, Florian, Farzad Saidi, and Glenn Schepens (2019) "Life below Zero: Bank Lending under Negative Policy Rates," *Review of Financial Studies*, 32 (10), 3728–3761.

John, Kose, Anthony Saunders, and Lemma W. Senbet (2000) "A Theory of Bank Regulation and Management Compensation," *Review of Financial Studies*, 13 (1), 95–125.

Kanatas, George (1986) "Deposit Insurance and the Discount Window: Pricing under Asymmetric Information," *Journal of Finance*, 41 (2), 437–450.

Kandrac, John and Bernd Schlusche (2021) "Quantitative Easing and Bank Risk Taking: Evidence from Lending," *Journal of Money, Credit, and Banking*, 53 (4), 635–676.

Keating, Thomas and Marco Macchiavelli (2017) "Interest on Reserves and Arbitrage in Post-Crisis Money Markets," Finance and Economics Discussion Series 2017-124. Washington: Board of Governors of the Federal Reserve System.

Keeley, Michael C. (1990) "Deposit Insurance, Risk, and Market Power in Banking," *American Economic Review*, 80 (5), 1183–1200.

Kim, Kyungmin, Antoine Martin, and Ed Nosal (2020) "Can the U.S. Interbank Market Be Revived?" *Journal of Money, Credit, and Banking*, 52 (7), 1645–1689.

Klee, Elizabeth, Ed Nosal, and Jacob Faber (2019) "Short Rates and a Large Balance Sheet: Exploring Costs and Competition in Money Markets," working paper.

Kreicher, Lawrence L., Robert N. McCauley, and Patrick McGuire (2013) "The 2011 FDIC Assessment on Banks' Managed Liabilities: Interest Rate and Balance-Sheet Responses," BIS working paper 413.

Landais, Camille (2015) "Assessing the Welfare Effects of Unemployment Benefits Using the Regfression Kink Design," *American Economic Journal: Economic Policy*, 7 (4), 243–278.

Martin, Antoine, James McAndrews, Ali Palida, and David Skeie (2013) "Federal Reserve Tools for Managing Rates and Reserves," Federal Reserve Bank of New York Staff Reports, No. 642.

Pennacchi, George (2006) "Deposit Insurance, Bank Regulation, and Financial System Risks," *Journal of Monetary Economics*, 53 (1), 1–30.

Quarles, Randal K. (2020) "The Economic Outlook, Monetary Policy, and the Demand for Reserves," Speech at the Money Marketeers of New York University, New York, NY.

Ronn, Ehud I. and Avinash K. Verma (1986) "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model," *Journal of Finance*, 41 (4), 871–895.

Schulhofer-Wohl, Sam and James Clouse (2018) "A Sequential Bargaining Model of the Fed Funds Market with Excess Reserves," Federal Reserve Bank of Chicago working paper, 2018-08.

Stein, Jeremy C. (2013) "Overheating in Credit Markets: Origins, Measurement, and Policy Responses," Speech at the "Restoring Household Financial Stability after the Great Recession: Why Household Balance Sheets Matter" research symposium sponsored by the Federal Reserve Bank of St. Louis, St. Louis, Missouri.