THE UNIVERSITY OF CHICAGO

ESSAYS IN FINANCIAL ECONOMICS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
AND
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
JINGTAO ZHENG

CHICAGO, ILLINOIS
JUNE 2024

To Ying Zheng, Wenzhu Chen, and Jingyi Tang.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am sincerely grateful to my dissertation committee, Raghuram Rajan, Wenxin Du, Ralph Koijen, Stefan Nagel, and Lars Peter Hansen, for their invaluable guidance, insightful advice, and strong support at each stage of my pursuit of the PhD. They not only taught me the tools of research but, more importantly, showed me what good research should look like. Their mentorship has profoundly shaped my research interests and steered me toward becoming an independent thinker. I am deeply grateful for their long-lasting impact on my professional development.

I also want to thank Doug Diamond, Niels Gormsen, Zhiguo He, Mark Mitchell, Lubos Pastor, Amir Sufi, Quentin Vandeweyer, Robert Vishny, and other Booth faculty for the stimulating discussions and helpful comments.

A special thanks goes to Jennifer Huang for introducing me to the world of finance research and setting me on this exciting journey. Her help and encouragement have been truly invaluable.

In addition, I would like to thank Mihir Gandhi, Pengyu Ren, Xiao Zhang, Ben Marrow, and other Booth PhDs for their support during this journey.

I am truly appreciative of the work done by the Booth PhD office. Malaina Brown, Cynthia Hillman, Kim Mayer, Kelly McDonell, Logan Ward, and Raven Davis have been very supportive in navigating the entire process.

Finally, I would like to express my sincere gratitude to my family - my father Ying Zheng, my mother Wenzhu Chen, and my wife Jingyi Tang. Their constant support and unconditional love have been my backbone. I couldn't have come this far without them.

ABSTRACT

This dissertation, consisting of 2 chapters, explores how institutional investors play an important role in financial markets.

In Chapter 1, I first present a new asset pricing anomaly: a simple dividend-based currency strategy, which shorts a currency on the date its country's recent aggregate dividend payment by listed companies is large, exhibits a significant Sharpe ratio and alpha not explained by standard factors in the currency market. To understand this anomaly, I identify the significant price impact of predetermined dividend payments on exchange rates around payment dates. I propose a dividend repatriation channel where benchmark investors (ETFs and mutual funds) predictably repatriate a certain proportion of dividends received in local currency. I build a model in which heterogeneous financial intermediaries with limited risk-bearing capacity accommodate benchmark investors' currency demands stemming from dividend repatriation flows. In line with the model's implications, I find that the price impact of dividend flows on FX around the payment date is large when the intermediary capital ratio is low, CIP deviations are large, and FX implied volatilities are high. My findings have implications for currency-market elasticity, capital regulations, and FX regimes.

In Chapter 2, I develop a machine learning procedure to estimate investors' demand system in high dimension, which accommodates a large universe of stock characteristics, including price-based characteristics (e.g., momentum, valuation ratio, etc.). I propose an identification strategy based on the inter-temporal structure of latent demand to address the endogeneity of price-based characteristics, in addition to the instrumental variables. Using the U.S. stock market data, I illustrate how we can use the estimated high-dimensional demand system to analyze the time variations in the importance of stock characteristics in investors' holdings, each stock characteristic's impact on cross-sectional stock returns, and identify which investors are significant for characteristic pricing.

CHAPTER 1

DIVIDEND FLOWS AND THE FOREIGN EXCHANGE RATE

1.1 Introduction

How do capital flows impact the foreign exchange rate (FX)? This is a central question in international finance. The answer to this question reflects how the currency market functions, especially the interaction between the demand and supply in the currency market. The previous literature emphasizes capital flows' information content and how the information is incorporated into exchange rates (Evans and Lyons 2002, Lyons 2001). Recent developments highlight the key roles played by financial intermediaries with limited risk-bearing capacity in segmented capital markets (Camanho et al. 2022, Gabaix and Maggiori 2015, Itskhoki and Mukhin 2021).

My paper provides new insights into this question by examining the FX dynamics related to dividend flows to foreign investors – a specific type of capital flow that is recurring, predictable, and informationless on the payment date. Dividend payments are predetermined: at the company level, all dividend information is released on the dividend announcement date, including the dividend amount and other dividend-related dates. Aggregated to the currency level, dividend payments are informationless on the payment dates. Standard asset pricing models imply the effects of flows on asset prices should be mainly on the announcement dates, while the effects on the actual realization dates should be small. Surprisingly, as my paper will show, the payment date effect of dividends on exchange rates is significant, while the anticipation effect before the payment date is limited, and the announcement date effect is negligible. Although dividends have been used in identifying price impact in domestic stock markets (Hartzmark and Solomon 2022, Schmickler 2022), they have not been studied by international economists. My paper fills this gap.

Specifically, I present new facts on how dividend flows affect FX dynamics among G10 currencies. G10 currencies are ten of the most liquid and most traded currencies: Australian dollar (AUD), Canadian dollar (CAD), Euro (EUR),[1] Japanese yen (JPY), New Zealand dollar (NZD), Norwegian krone (NOK), British pound (GBP), Swedish krona (SEK), Swiss franc (CHF), United States dollar (USD).[2] G10 is an ideal empirical setting for identifying the price impact of dividend flows on FX for four reasons. First, G10 countries have large stock markets. Over the sample period from 2001 to 2022, the average stock-market-capitalization-to-GDP ratio ranges from 0.36 for New Zealand to 2.14 for Switzerland. Second, other countries' ownership of each G10 country's stock market is substantial. The sample average foreign ownership ranges from 17.6% in the United States to 60% in Switzerland. Third, G10 currencies have fewer confounding central bank direct interventions in the FX market than emerging market currencies. Fourth, G10 currencies are the most liquid currencies. In a counterfactual world without central bank interventions, the price impact of dividend flows on G10 currencies should be smaller than other currencies. In this sense, I interpret my estimates as a lower bound for the price impact for other currencies.

As G10 currencies are the most researched and traded currencies by market participants, one might expect predictable flows based on predetermined dividend payments to have negligible effects. However, this is not the case. As motivating evidence, I present a simple dividend-based currency strategy, which shorts a currency if its country's recent aggregate dividend payment is large, and closes the position the next day. The strategy aims to capture the local currency's depreciation pressure shortly after its dividend payment. This strategy can be implemented in real time, as dividend payments – both dates and amounts – are

1. The euro area (aka. eurozone) consists of 19 countries that use the Euro: Belgium, Germany, Ireland, Spain, France, Italy, Luxembourg, the Netherlands, Austria, Portugal, Finland, Greece, Slovenia, Cyprus, Malta, Slovakia, Estonia, Latvia and Lithuania. Starting January 1, 2023, Croatia became the 20th member of the eurozone.

2. See Table 1.10 for the abbreviations of The G10 currencies and their countries/currency areas used in the paper. For the definition of G10 currencies, please refer to in Article I(2) in `https://www.occ.gov/news-issuances/news-releases/2014/nr-occ-2014-157e.pdf`. An alternative definition of G10 includes Danish krone (DKK). I do not include DKK as it is always pegged to EUR for my sample period.

known beforehand. Surprisingly, I show this simple strategy has a significant Sharpe ratio and alpha that are not explained by standard factors in the currency market, including the dollar, carry, momentum, and value factors. The results are robust under different parameters and reasonable assumptions about transaction costs faced by institutional investors.

To understand this asset pricing puzzle, I identify empirically the magnitude of the price impact of dividends on the foreign exchange rate. My identification strategy exploits the fact that dividends are predetermined and hence informationless on payment dates. Therefore, these dividend payments should not contain contemporaneous information that affects the foreign exchange rate after its announcement, specifically around the payment date. I focus on the payment date on which the dividends are large, as their effects on foreign exchange rates should be the most prominent. The baseline panel regression includes controls, time fixed effects, and currency fixed effects. The control variables, i.e., stock market returns and FX implied volatilities, serve to account for alternative channels, such as the portfolio rebalancing channel described in Camanho et al. (2022), where global equity investors adjust their portfolio allocations in response to stock market returns, and such rebalancing is more intense under higher FX volatility. The time fixed effect, at the date level, addresses FX seasonality, and the month-end/quarter-end effect due to month-end/quarter-end rebalancing. The currency fixed effect controls for currency-specific trends throughout the sample.

Regressing the cumulative change of the foreign exchange rate on the large dividend indicator reveals a consistent pattern: upon and after its large dividend payment, the local currency depreciates against USD. Two days after the dividend payment date, the cumulative currency depreciation against USD is around 4.70 basis points. Eight days after the dividend payment date, the local currency has depreciated to 6.48 basis points, and it shows signs of slight reversion afterward. In contrast, the price effect before the payment date (aka. anticipation effect) is limited, even though the dividend payment is known and imminent. Moreover, the FX effects of dividends around the dividend announcement date are economically small and statistically insignificant. These empirical findings are robust under various identification strategies. However, they are in sharp contrast to the predictions of standard asset pricing models. First, as asset prices should only respond to new information, standard models suggest the effects should be largest on the announcement date. Second, standard models also imply the effects should be small around the payment date, otherwise, this implies (risky) arbitrage opportunity. Third, as forward-looking speculators should have pre-positioned in advance, we should also see FX movement before the dividend payment actually happens. Nevertheless, none of these predictions by standard models are supported by my empirical findings.

I build a model of currency demand and supply that explains the FX dynamics of exchange rates around dividend payment dates and dividend announcement dates. On one side, for dividends to move exchange rates shortly after the payment dates, a certain proportion of dividends must be repatriated and converted into other currencies. I propose a dividend repatriation channel, in which benchmark investors of global equities predictably repatriate dividends received in local currency shortly after receiving them. Benchmark investors include passive ETFs and mutual funds (i.e., index funds), which aim to track the performance (especially the total return) of their benchmark indices as closely as possible. More broadly, benchmark investors also include active funds practicing closet indexing (Cremers and Petajisto 2009, Cremers et al. 2016). Benchmark investors have been playing an increasingly important role in the global equity market. As Figure 1.1 shows, the market value of US-domiciled ETFs' foreign holdings as a percentage of the other G10 countries' stock market capitalization has grown from 0.7% in 2011 to 3.2% in 2020, more than quadruple in 9 years. In addition, US-domiciled mutual funds have grown from 1.93% in 2002 to 4.6% in 2011 to 6.6% in 2020.

Benchmark investors have particular incentives to repatriate dividends shortly after receiving them to minimize deviations from their benchmark equity indices, i.e., the tracking errors. Here is the reason: the index methodologies of mainstream equity indices assume the reinvestment of cash dividends into the index itself pro rata on the ex-date. As a concrete example, suppose an index has a 20% allocation in pound-denominated stocks and a 80% allocation in non-pound-denominated stocks. On the ex-date of cash dividends in pounds, the index will retain only 20% of the cash dividends in pounds and reinvest them into pound-denominated stocks. The remaining 80% of cash dividends will be converted into other currencies and reinvested into non-pound-denominated stocks. Although equity indices prescribe dividends to be reinvested on the ex-date, benchmark investors only receive dividends on the payment date, which lags behind the ex-date. Therefore, benchmark investors have incentives to act quickly, as further delay may lead to increased tracking errors. Regarding the exact implementation of reinvestment, the fund manager can either repatriate to other currencies and reinvest directly into the underlying stocks or, more commonly, repatriate the dividends back to the fund's home currency and use futures to establish effective exposures, which is more cost-effective. In either case, a certain proportion of the dividends are predictably repatriated out of the currency that pays the dividends and converted into other currencies. Using detailed daily positions of ETFs, especially cash positions in different currencies, I present a case study that provides empirical evidence for the dividend repatriation channel.

On the other side of the currency market are financial intermediaries, including banks, dealers, and arbitrage capital like hedge funds and proprietary desks. Because the intermediaries have limited risk-bearing capacity, they need to be compensated to accommodate the currency demand from benchmark investors. Moreover, intermediaries have different levels of sophistication in parsing the FX implications of dividend payments, which results in different beliefs about future exchange rates. Some intermediaries (e.g., speculators) are attentive to dividend payments and have rational expectations of future exchange rates. Other intermediaries (e.g., uninformed liquidity providers) are less sophisticated. They do not understand the implications of dividend payments on exchange rates. Their expectation of the next period's exchange rate is always the long-run equilibrium exchange. With the presence of uninformed liquidity providers, predictable dividend flows will have a significant payment date effect, because for these intermediaries the dividend payments are as if they are unexpected, despite being public information before the payment dates. The speculators cannot correct all the mispricing because deploying their limited capital to conduct risky arbitrage is costly. In equilibrium, they do not aggressively take short positions far in advance. With reasonable calibration of the proportion of sophisticated vs unsophisticated intermediaries, the model quantitatively explains the large payment date effect, the limited anticipation effect, and the negligible announcement date effect of dividends.

The model has further implications for the time variation of the price impact of dividends on exchange rates around payment dates: during the periods when financial intermediaries' risk-bearing capacity is lower, the price impact of dividend flows should be larger. In the model, intermediaries' risk-bearing capacity depends on their balance sheet constraints and the currency market volatilities. I use the intermediary capital ratio and the CIP deviations to proxy for the balance sheet constraints, as the intermediary capital ratio is the cause while the CIP deviations are the result. I use the currency implied volatilities to proxy for the FX market volatility, as the implied volatility is the forward-looking measure. Consistent with the model's predictions, I find that the price impact of dividend flows on FX around the payment date is large when the intermediary capital ratio is low, CIP deviations are large, and FX implied volatilities are high.

I conclude by discussing the implications of my findings on currency market elasticity, capital regulations, and FX regimes. A back-of-envelope calculation shows \$8.1 billion US dollars moves G10 against USD by 1%. At first glance, this falls in the ballpark of existing estimates in the literature and is consistent with the recent literature on the inelastic market hypothesis pioneered by Gabaix and Koijen (2021). However, the fact that dividend flows move the foreign exchange around the payment dates is more puzzling, as the model in Gabaix and Koijen (2021) predicts that most of the price effect should happen on the announcement date while the price effect on the payment date should be small if agents are forward-looking. My estimates also suggest one standard deviation (3.1%) decrease from the mean (7.38%) of the intermediary capital ratio is associated with a price impact twice as large. I also find evidence suggesting the price impact of dividends on FX is larger in the freely floating regime compared to other regimes.

The paper is structured as follows: after a brief literature review, Section 1.2 introduces the datasets I use in the empirical analysis. As background knowledge, Section 1.3 presents the stylized facts on dividends in the G10 countries. Serving as motivating evidence, Section 1.4 presents a dividend-based currency strategy that has a significant Sharpe ratio and alpha not explained by standard FX factors. To understand this anomaly, Section 1.5 identifies the FX dynamics around the payment dates, in addition to the announcement date effect. Section 1.6 analyzes the underlying mechanism and proposes the dividend repatriation channel. It presents a model that explains the significant payment effect, the limited anticipation effect, and the negligible announcement effect of dividends. Consistent with the additional model implications, Section 1.7 shows that the price impact of dividend flows on FX around the payment date is large when the intermediary risk-bearing capacity is low. Section 1.8 discusses the implications of my estimates for FX elasticity, capital regulations, and FX regimes. Section 1.9 concludes with a discussion with open questions for future research.

1.1.1 Related Literature

My paper is related to three strands of literature. First, my paper is related to the literature on capital flows and their impact on the foreign exchange rate. Maggiori (2022) provides a comprehensive review of the literature. Theoretically, Evans and Lyons (2002) present an exchange rate model highlighting the information content of order flows. Gabaix and Maggiori (2015) provides a theory of foreign exchange determination in which capital flows drive exchange rates by altering the balance sheets of intermediaries with limited risk-bearing capacity. Itskhoki and Mukhin (2021) show that financial shocks (i.e., noise-trader demand shock) are the only plausible shocks to explain exchange rate dynamics. Hau and Rey (2006) and Camanho et al. (2022) develop equilibrium models in which exchange rates, stock prices, and capital flows are jointly determined. They highlight the portfolio rebalancing channel of global equity investors. In contrast, my paper highlights informationless dividend flows impact FX shortly around payment dates, due to the dividend repatriation channel. Empirically, as capital flows are likely to be endogenous to exchange rates and financial conditions, most papers estimate the price impact of capital flows using one-off events and focus on the announcement date effect. Hau et al. (2010) use the redefinition of the MSCI Global Equity Index in 2001 and 2002, a switch of index weights from market capitalization to freely floating. They find countries with a relatively increasing equity representation have a relative currency appreciation on the announcement date of the index change. Broner et al. (2021) use the unexpected announcement of index inclusion into local-currency sovereign debt indexes of Citigroup WGBI and JP Morgan GBI-EM, and find that index-inclusion-induced inflow leads to an appreciation of the country's currency in the two days following the announcement. However, they find no effect during the implementation period between 2 and 6 months after the announcement date. In contrast, Raddatz et al. (2017) find that large benchmark changes (such as upgrades and downgrades of countries) are associated with abnormal returns in asset prices and exchange rates around those events, both on the announcement and effective dates of these changes. Some other papers use more frequent events to estimate the price impact. Camanho et al. (2022) apply the granular instrumental variable (GIV) approach to funds' rebalancing flows. Aldunate et al. (2022) use Chilean pension funds flows induced by a Chilean financial advisor' market timing recommendations. In terms of the nature of flows, the closest paper to mine is Pandolfi and Williams (2019), which uses mechanical rebalancings induced by the J.P. Morgan Government Bond Index–Emerging Markets Global Diversified (GBI-EM Global Diversified) 10% index weight cap of any single country. This feature may not be widely recognized compared to dividend payments, the latter of which are closely watched by market participants.[3] In addition to the reduced form approach, Koijen and Yogo (2024) propose a structural form approach based on a demand system of global investors.

Second, my paper is related to recent developments investigating the relationship between flows and prices, primarily in the stock markets. Gabaix and Koijen (2021) develop a theory of inelastic demand under rigid institutional investors' mandate and estimate the price elasticity of aggregate stock market demand using GIV. In their model, the largest effect happens upon news of flows, not when flows actually happen. Closely related to my paper is Hartzmark and Solomon (2022). They study the effect of dividends on the aggregate equity market. Despite being predetermined, dividends move the stock market due to the reinvestment channel. Schmickler (2022) finds that dividends generate payment date price pressure for peer stocks in the portfolio, but not on the announcement date. In contrast, I propose that dividend flows move the foreign exchange rate due to a different mechanism, i.e., the dividend repatriation channel. The existing literature related to dividend repatriation mostly focuses on corporate shareholders' repatriation of foreign subsidiaries' dividends, especially when there is a repatriation tax change or one-time tax holiday (De Simone et al. 2019, Hanlon et al. 2015). This paper instead emphasizes the role of foreign portfolio investors, particularly benchmark investors such as ETFs and mutual funds, in regular dividend repatriation.[4]

Third, my paper is related to the literature on intermediary asset pricing, where financial

3. Analysts at banks regularly distribute dividend information to their clients, e.g., hedge funds.

4. Corporate shareholders' ownership of foreign subsidiaries is counted as direct investment rather than portfolio investment, according to the Balance of Payments.

intermediaries with limited risk-bearing capacity play a key role in FX determination. Theoretically, He and Krishnamurthy (2013) propose a model where the marginal investor is a financial intermediary. Empirically, He et al. (2017) find that shocks to the equity capital ratio of financial intermediaries have significant explanatory power for cross-sectional variation in expected returns in many asset classes, including currencies. Reitz and Umlandt (2021) further refine the intermediary capital ratio for the currency markets using the balance sheet data of the top three foreign exchange dealers. Du et al. (2018) finds that banks' balance sheet constraints have a causal effect on asset prices, as reflected in deviations from the covered interest rate parity condition (CIP). Interpreted more generally, financial intermediaries also include arbitrage capital like proprietary desks, macro hedge funds, active investment managers, etc. Their limited risk-bearing capacity leads to limits of arbitrage, pioneered by De Long et al. (1990), Shleifer and Vishny (1997), and Gromb and Vayanos (2002).

1.2 Data

The dividend information is from Compustat Global and the Center for Research in Security Prices (CRSP).[5] For countries other than the USA, I use Compustat Global, while I use CRSP for dividend information in the USA. The dividend information includes dividend size, announcement date, ex-date, and payment date. On a few occasions when dividend payment dates coincide with weekends, I assume the dividends occur on the following business day. I focus on cash dividends and keep common/ordinary shares. For stocks with dual-listing or multiple currencies, I use their primary listing information.

The G10 currency market is a 24-hour market. In contrast, stock markets in each country have operating hours locally, and databases like Compustat Global and CRSP record date information in their respective time zones. The cutoff time in the standard sources of the

5. Omitted dividends may not be recorded in either database. However, this does not affect the dividend-based currency strategy or the identification, as the decision to skip a dividend is announced before the payment date.

foreign exchange rates may not necessarily align with the local stock market closing time. e.g., WM/Refinitiv FX Benchmark Rates have the cut-off time at London 4 p.m., while Bloomberg provides three pre-fixed cut-off times.[6] Misalignment of FX cut-off time and stock market closing time may lead to asynchronicity issues, especially in the daily frequency analysis.

To alleviate the concern of asynchronicity, I assemble a novel dataset of daily changes in foreign exchange rates of each currency, aligned with each country's local stock market closing time. To do so, I use the hourly spot exchange rate from WM/Refinitiv intraday fixing and snapshot the exchange rates at the closest hour to the stock market closing time, as Table 1.4 shows. The WMR Intraday Spot Rate service was launched in 2001. It provides hourly spot rates from Monday 06:00 in Hong Kong/Singapore until Friday 22:00 in the UK.[7] The foreign exchange rates are quoted against US dollars using market conventions.[8] In my analysis and throughout the paper, I express all exchange rates in units of USD (or a basket of currencies) per local currency. Therefore, a negative change means the local currency depreciates against USD. The sample period is from January 2001 to June 2023.

I construct three measures of FX change: against USD, against a value-weighted G10 basket, and against an equal-weighted G10 basket. In the value-weighted G10 basket for currency i, the weight of the currency pair j/i is proportional to the foreign country j's ownership of the stock market of i, proxied by data from Coordinated Portfolio Investment Survey (CPIS).

The ETF daily positions are from ETF Global, which covers ETFs listed in the US. Starting from April 2017, ETF Global Data is primarily sourced directly from fund sponsors,

6. BGN closes 5 p.m. Friday EST (New York cut), BGNL closes at London 6 p.m. (London cut) and BGNT closes at Tokyo 8 p.m. (Tokyo cut). Both London and Tokyo cut close at 5 p.m. EST on Friday. Some emerging market currencies have their cut-time limited to when the local market closes.

7. For more details, see `https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/wm-refinitiv-methodology.pdf`

8. That is, in units of local currency per USD, except for EUR, GBP, AUD, NZD.

custodians, distributors, and administrators. [9] Its *Constituents* file contains actual holdings of the many ETFs at daily frequency, including cash, derivatives, and underlying. I use Morningstar for longer time series of ETF and mutual fund quarterly holdings.

Information on cross-border flows and positions is from Balance of Payments and International Investment Position, downloaded from the International Monetary Fund (IMF). In addition to standard items related to trade and current account surplus, I focus on items related to portfolio investments[10] of different countries. It reports the dollar value of a country's ownership in other countries' assets (e.g., equity and debt securities), and the dollar value of a country's assets being owned by other countries. In the financial account, flows such as net acquisition of financial assets (i.e., the purchase of other countries' assets) and net incurrence of liabilities (i.e., assets being purchased by other countries) are reported. In the capital account, investment income from portfolio investment, including dividends and interest, is reported. The bilateral ownership information of portfolio investment comes from the Coordinated Portfolio Investment Survey (CPIS) by the IMF.

1.3 Stylized Facts about Dividends

In this section, I present stylized facts about dividends in G10 currency areas, showing that they are predetermined, substantial, and concentrated. In this paper, G10 countries/currency areas refer to major countries that use G10 currencies. They are: Australia (AUS), Canada (CAN), Switzerland (CHE), Euro area (EUR), United Kingdom (GBR), Japan (JPN), Norway (NOR), New Zealand (NZL), Sweden (SWE), United States (USA).

Dividends are predetermined. At the company level, there are four important dates related to dividends: the announcement date, the ex-date, the record date, and the pay-

9. `https://wrds-www.wharton.upenn.edu/documents/1719/ETF_Global_Data_Package_-_U.S._Listed_-_2021_-_1.1.21.pdf`

10. Portfolio investment is defined as cross-border transactions and positions involving debt or equity securities, other than those included in direct investment or reserve assets. See Sixth Edition of the IMF's Balance of Payments and International Investment Position Manual (BPM6).

ment date. The announcement date is the date when a company announces its dividend information, including dividend amount and other dividend-related dates. The ex-date is the date on and after which shareholders who buy the stock will not receive a dividend. The record date is the date on which registered shareholders in the company's book will be entitled to receive dividends.[11] The payment date is the date when the dividend is actually paid to shareholders. I aggregate the companies' dividends to country/currency area level by payment date.

All dividend information is revealed on the dividend announcement date,[12] including dividend amount and other dividend-related dates, in all G10 countries/currency areas except Japan. For Japan, companies typically do not confirm the exact dividend amount before the ex-date, though the dividend guidance is usually available almost one year in advance. In any case, on the actual payment date, the dividend is informationless. Table 1.2 shows the calendar days between the announcement date and the payment date for countries except Japan, and calendar days between the ex-date and the payment date for Japan. There is a big time gap - the average lead time is 58 days, with a median of 55 days. Such a time lag should be enough for the market to digest the information released on the announcement date.

Dividends are substantial. With aggregate dividend yields ranging from 2% to 5%, large stock market in G10 countries implies large aggregate dividend payments. Indeed, Table 1.1 shows the stock-market-capitalization-to-GDP ratio ranges from 0.36 in New Zealand to 2.14 in Switzerland over 2001 to 2022 sample period, while the dividend-to-GDP ratio ranges from 1.5% in Euro area to 3.7% in Australia.

Importantly, due to large foreign ownership (Table 1.1), dividends paid to foreign investors can be substantial. In fact, data from the Balance of Payments reveals this pattern.

11. Depending on the settlement cycle, the ex-date is typically one day before the record date.

12. Even in rare circumstances where a company needs to skip a dividend payment, it will announce this decision on the announcement date.

Table 1.3 summarizes the dividends paid to foreign portfolio investors and the dividends received from foreign portfolio investments. In BOP, these items are recorded as primary income. See Appendix 1.10.1 for detailed indicators. On an annual basis, dividends paid to foreign investors are comparable to portfolio investment flows, either in equity and debt. Specifically, the average dividends paid to foreign investors is $36.7 billion per year, while the average purchase of foreign equity is $37.1 billion and the average purchase of foreign debt is $64 billion. Compared with trade flows, dividends paid to foreign investors are also of the same order of magnitude.

Dividends are concentrated. Dividend payments are not evenly distributed throughout the years. As Figure 1.2 shows, dividends can be concentrated in some days, weeks, and months. e.g., the top 5% largest dividend payment dates contribute to a significant proportion of the total dividend payment in a year, ranging from 28% in the United States to more than 60% in Japan. When calculating how many days of the largest dividend payments contribute to more than 50% of total dividends within a currency-year, the number ranges from 2.5 days in Switzerland, to 18.5 days in the Euro Area, and up to 30.5 days in the US. Zooming out to the monthly level, dividend payments in the United States are concentrated in the last month of each quarter (March, June, September, December), while in the Euro area, they are concentrated in May. In Japan, dividends are concentrated in June and December.

There are several reasons for the concentration of dividends. First, due to traditions and customs in a country, companies may follow a similar fiscal-year calendar. For example, in Japan, most companies have a fiscal year-end on March 31, following the government fiscal year calendar. The similarity in corporate fiscal calendars leads to the concentration of dividend dates. Second, bigger companies pay larger dividends. With the company size being skewed, the dividends may be dominated by a few large companies. [13]

13. For example, Taiwan Semiconductor Manufacturing Company Limited (TSMC) is the largest company primarily listed in the Taiwan Stock Exchange. As of 2022 year-end, its market capitalization is 379

1.4 Dividend-Based Currency Strategy

In this section, I present a dividend-based currency strategy on G10 currencies. The strategy takes the following simple format: sell the currency if the country has large dividends in the past few days against USD and hold the position for one day. This strategy has a significant Sharpe ratio and alpha not explained by standard FX factors, including carry, dollar, momentum, and value factors, despite it only uses publicly available dividend payment information known ex-ante.

The log excess return of selling currency i against USD, and holding the position for one day is:

$$rx_{t+1}^{k} = f_t^k - s_{t+1}^k \approx -\Delta s_{t+1}^k + (i_t^{US} - i_t^k)$$

where s_t^k and f_t^k are log spot exchange rate and log 1-day forward exchange rate of currency k respectively, in terms of units of USD per local currency, i.e., currency k is the base currency. i_t^{US}, i_t^k is 1-day risk-free rate in the USA and country k, respectively. As my sample of WMR intraday hourly fixing does not contain 1-day forward exchange rate, I use $f_t^k \approx s_t^k + i_t^{US} - i_t^k$ to approximate it, where the risk-free rates are (annualized) 3-month risk-free rates divided by 365.

With transaction costs, the log excess return of selling currency k against USD, and holding the position for one day is:

$$rx_{t+1}^{k} = f_t^{k,b} - s_{t+1}^{k,a} \approx -\Delta s_{t+1}^k + (i_t^{US} - i_t^k) - TC$$

where the transaction cost (TC) is the bid-ask spread of spot exchange rates. The bid–ask spreads from WMR are based on indicative quotes. They are too large compared to actual effective spreads in FX markets (see, e.g., Lyons 2001, Menkhoff et al. 2012). As G10

billion New Taiwan dollars, 24% of the total stock market capitalization. Its quarterly dividend payments throughout 2022 sum up to 285 billion New Taiwan dollars, around 18% of the total dividend payment in the Taiwan stock market.

currencies are the most liquid currencies in FX markets, they have very tight bid-ask spreads for institutional investors, mostly a fraction of 1 basis point. Moreover, large intermediaries may collect the bid-ask spread when trading with clients or trading at close to the mid-price in interdealer markets. Therefore, I assume a constant 1 basis point bid-ask spreads for G10 currencies in the below discussion.

The dividend-based currency strategy takes the following form: for each country/currency area k and date t, if in the previous l days, the combined dividend payments in the country k rank in its top p-percentile in the rolling 1-year window, then we sell currency k against USD, and hold the position for one day. If there are several currencies that satisfy this criterion, the strategy puts \$1 on each position.

Figure 1.3 shows the cumulative excess return of the dividend-based currency strategy in percentage points, for the parameters $l = 2, p = 5\%$. In other words, the strategy sells a currency against USD if the combined dividend payments in the previous 2 days rank in the top 5% percentile in the rolling 1-year window of that country. Over the sample period from 2001 to 2023, the strategy earns 4.4% return annually before the transaction cost (blue line), with 0.68 Sharpe ratio, despite it only takes positions on 25% of the trading days. After the transaction cost (orange line), the annualized return is 3.6% with Sharpe ratio being 0.56. Noticeably, the performance of the trading strategy is better after the Global Financial Crisis (GFC) than before.

The top half of Table 1.5 shows the results are robust across different parameters of the lookback window l. Both the annualized returns and Sharpe ratio are statistically significant. Note that the strategy achieves this annualized returns by taking FX positions on $\approx 30\%$ days of trading days. The standard errors of the Sharpe ratio are calculated using the formula in Lo (2002).

The bottom half of Table 1.5 further demonstrates that the dividend-based currency strategy generates alpha not explained by standard factors in the currency market. To show this, I run the following factor-spanning regression (Fama and French 2018) at the monthly frequency:

$$rx_t = \alpha + \beta_{DOL} DOL_t + \beta_{CAR} CAR_t + \beta_{MOM} MOM_t + \beta_{VAL} VAL_t + \epsilon_t \tag{1.1}$$

The rx_t are log excess returns of the dividend-based currency strategy aggregated to the monthly frequency. The dollar factor DOL_t is from Verdelhan (2018). The carry factor CAR_t is from Lustig et al. (2011). The momentum factor MOM_t is from Menkhoff et al. (2012). The value factor VAL_t is from Asness et al. (2013). As expected, the strategy has a significant loading on DOL_t, since it sells a currency against USD. The strategy's loadings on other factors are economically small and statistically insignificant. Importantly, the alpha is economically large and statistically significant. The monthly alpha is around 30bp. When annualized, the alpha accounts for almost all the annualized returns of the strategy.

1.5 Identification of the Price Impact of Dividend Flows

To understand the anomaly, in this section, I identify the price impact of dividends on exchange rates around the payment dates. I show that the local currency depreciates shortly after its country's large dividend payments. In contrast, the anticipation effect before the payment date is limited. In addition, I show that the dividend announcement effect on exchange rates is insignificant.

1.5.1 Around Dividend Payment Dates

My identification strategy exploits the fact that dividends are predetermined hence informationless on payment dates. In fact, companies make dividend decisions using information up to their announcement dates. Therefore, dividends do not contain any contemporaneous information that affects exchange rates after its announcement, specifically around the payment date.

I focus on the payment date on which the dividends are large, as their effects on foreign exchange rates should be the most prominent. Specifically, let $\mathbb{D}_{k,t}$ be the indicator for large dividends, i.e., $\mathbb{D}_{k,t}$ equals 1 if country/currency area k has a large dividend payment on date t. For concreteness, I define a large dividend as being among the top 5% largest within the currency-year on the payment date t.

The baseline panel regression is as follows:

$$\ln E_{k,t+h}^{US/LC} - \ln E_{k,t-1}^{US/LC} = \alpha_h + \beta_h \mathbb{D}_{k,t} + Controls + \gamma_k^{(h)} + \xi_t^{(h)} + \epsilon_{k,t}^{(h)}, \quad h = -10, ..., 0, ..., 10 \tag{1.2}$$

The left-hand side is the cumulative log change of the exchange rate of currency i against USD[14] in basis points from date $t-1$ to $t+h$, with FX cut-off time aligned with the local stock market closing time. $t = -1$ is one day before the payment date, which I normalize the cumulative change to be 0. The parameters of interest are β_h's on the dividend indicator $\mathbb{D}_{k,t}$. The *Controls* include local stock market returns and FX implied volatilities. For stock market returns, I use daily changes in each country's primary stock index.[15] For FX implied volatilities, I use the 6-month at-the-money (ATM) implied volatility for each currency against the USD.[16] The control variables serve to account for alternative channels, such as the portfolio rebalancing channel described in Camanho et al. (2022), where global equity investors adjust their portfolio allocations in response to stock market returns, and such rebalancing is more intense under higher FX volatility. The time fixed

14. The results are similar if using cumulative log change against a value-weighted and an equal-weighted G10 basket.

15. Specifically, S&P/ASX 200 Index (Australia), S&P/TSX Composite Index (Canada), Swiss Market Index (Switzerland), Euro Stoxx 50 (Euro area), FTSE 100 Index (United Kingdom), NIKKEI 225 (Japan), OBX STOCK Index (Norway), S&P/NZX 50 Index (New Zealand), OMX Stockholm 30 Index (Sweden), S&P 500 Index (United States).

16. The results using other tenors of implied volatility are almost exactly the same.

effect, at the date level, addresses FX seasonality[17] and the month-end/quarter-end effect due to month-end/quarter-end rebalancing. If the spillover effect of another country's large dividend payment is similar across currencies, then the time fixed effect accounts for this as well. The currency fixed effect controls for currency-specific trends throughout the sample. The standard errors are two-way clustered at the date level and the currency level, to take into account the potential correlation of residuals due to common factors and overlapping samples.

Table 1.6 compares the coefficients β_h estimated by the variants of Eq (1.2), which incrementally add controls and fixed effects. Panel *OLS* shows the estimates without any controls and fixed effects. Panel *OLS with Controls* shows the estimates controlling for stock market returns and FX implied volatilities. Panel *OLS with Controls and Time Fixed Effects* further controls for time fixed effects. Panel *Two-Way Fixed Effects with Controls* is the baseline regression results, which are plotted in Figure 1.4. Figure 1.11 further shows the comparison of coefficients under different specifications. All the specifications show a consistent pattern: upon and shortly after the dividend payment dates, the local currency depreciates against USD. Indeed, the cumulative currency depreciation against USD is 4.70 basis points two days after the dividend payment date. After eight days, the local currency depreciates against USD by 6.48 basis points. It shows signs of slight reversion afterward. In Section 1.6, I argue this depreciation pressure is due to the *dividend repatriation channel*, i.e., benchmark investors' predictable repatriation of dividends from the dividend currency to other currencies shortly after receiving the dividend payments.[18]

In contrast, the anticipation effect before the dividend event $t = 0$ is economically and statistically limited. The only statistically significant coefficient under the baseline specifi-

17. Fei (2023) documents the dollar depreciates by 54 basis points on average in the last 10 trade days of the calendar year and appreciates by 47 basis points in the first 10 trade days of the next year. Tse (2018) documents all the G10 currency futures yield negative returns in January and returns in April are positive.

18. Nevertheless, the response may be delayed in a few days as cash dividends may appear in an investor's account with a lag, as shown by the example in Section 1.6.2.

cation is at $t = -1$. As dividend payments are public information and ex-ante known, the anticipation effect may be due to some investors' pre-positioning by selling local currency in advance to take advantage of the benchmark investors' dividend repatriation. Alternatively, some investors may conduct the FX spot transaction 1 or 2 days before the dividend payment date, as the settlement date for FX spot transactions is T+2, i.e., two business days after the trade date.[19] When cash dividends in local currency appear on their cash account, they can directly use it to settle the FX spot transaction. Empirically, we see the anticipation effect is limited.

1.5.2 Around Dividend Announcement Dates

To study the FX dynamics around announcement dates of large dividends, I run a similar regression as in Eq (1.2), where t is the announcement date instead of the payment date. In other words, dividends at the company level are aggregated to the currency level by their announcement dates. $\mathbb{D}_{k,t}$ equals to 1 if country/currency k has a large dividend announcement on date t. For consistency with the payment date results, I define a dividend announcement as being large if it is among the top 5% largest within the currency-year. The *Controls* include local stock market returns and FX implied volatilities. As before, the time fixed effect and the currency fixed effect are included in the baseline regression, and the standard errors are two-way clustered at the date level and the currency level.

It is worth noting that dividends having the same payment date may not have the same announcement date, and vice versa. Therefore, when aggregating dividends to the currency level, there may not be a simple correspondence between large dividend announcement dates and large dividend payment dates. Nevertheless, as big companies contribute the most to the dividends, the set of companies on large dividend announcement dates are highly correlated with the set of companies on large dividend payment dates.

19. For USDCAD spot transactions, the settlement date is T+1, one business day after the trade date.

Figure 1.5 shows the announcement date effect of dividends on the foreign exchange rate. In contrast to the payment date effect, the announcement date effect is economically small and statistically insignificant. Specifically, from ten days before to ten days after a large dividend announcement date, all point estimates of price impact are between -2 to +2 basis points, and none of them is significant.

1.5.3 Robustness

The identification strategy in Section 1.5.1 only uses dividend size on the payment date, which is known ex-ante before the payment date. By using only the dividend size information ex-ante known, it extracts the predictable component of dividend flows. Importantly, the baseline identification does not use the actual dividend repatriation flows, which may contain contemporaneous information. The similarity of the empirical results between *OLS* and *Two-Way Fixed Effects with Controls* in Table 1.6 assures that the potential confounding variables should be unimportant.

Nevertheless, in the case of further identification concerns,[20] I develop alternative identification strategies in Appendix 1.10.4 to confirm that the pattern revealed in Section 1.5.1 is robust. The additional identification strategies include difference-in-difference (DiD) and synthetic controls, which can deal with the unspecified confounding variables in a more flexible way. As shown in Figure 1.13 and Figure 1.14, both methods confirm that dividends move the foreign exchange rate shortly after the payment dates.

1.6 Inspecting the Mechanism

In this section, I present a model that explains the FX dynamics of exchange rates around dividend payment dates and dividend announcement dates. On one side, I highlight the

20. One potential concern on the baseline regression Eq (1.2) is that it assumes the unspecified time-varying confounding variables affect all currencies in the same way, and hence can be absorbed by the time fixed effect.

currency demand from the dividend repatriation channel by benchmark investors, due to the treatment of cash dividends in the underlying mainstream equity index methodologies. On the other side, financial intermediaries with heterogeneous beliefs and limited risk-bearing capacity need to absorb the flows on their own balance sheets. Due to the time variation of their risk-bearing capacity, the price impact of dividend flows differs over time.

1.6.1 Treatment of Cash Dividends by Equity Indices

Equity index methodology pays particular attention to corporate actions. Related to my paper is its treatment of cash dividends. There are three kinds of returns associated with equity indices: price return, gross (total) return, and net (total) return. The price return is the change in the price index[21] level, which is the weighted average of the underlying price of constituents, without taking into account the regular cash dividends.[22] The gross return assumes the dividends are reinvested into the index itself. The net return further considers the dividend withholding tax for foreign investors, assuming the dividends are reinvested after the deduction of withholding tax. Importantly, equity indices do not have cash components. In the equity index calculation, the dividends are reinvested immediately on the ex-date.

The dividends are not only reinvested into the original stocks that pay the dividends. Instead, the dividends are reinvested to the entire portfolio pro rata.[23][24] Formally, denote the total amount of dividends in index points divided by the index level by α. On ex-date, each share count is scaled up by a factor of $1/(1-\alpha)$. See FTSE (2023) Section 4, MSCI

21. Some index providers like FTSE Russell use the terminology *capital return* and *capital index*.

22. A special cash dividend that is nonrecurring may affect the calculation of the price index.

23. Note that the index weight of the stock paying the dividend changes before and after its dividend ex-date, as the ex-dividend price is lower than the cum-dividend price.

24. For fund inflows/outflows, the proportional investing assumption is common in the literature of mutual funds like Lou (2012) and Chen (2024). Following this literature, Schmickler (2022) also assumes dividend payments are reinvested pro rata. Here, I emphasize the underlying reason, i.e., the specific treatment of cash dividends by the equity index methodology.

(2023) Section 2, and Appendix X.[25]

For the global equity indices, the underlying stocks are not denominated in a single currency. The treatment of cash dividends in the index calculation implies dividends will be repatriated abroad. For example, suppose an index has 20% allocation in the UK and 80% outside the UK. On the ex-date of a dividend paid by a UK company in GBP, the index calculation assumes that ≈80% of the dividend will be reinvested to stocks outside the UK (hence in currencies other than GBP), converted at the spot exchange rates on the ex-date, typically WMR London 4 p.m. fixing rate.

1.6.2 Dividend Repatriation by Benchmark Investors

Benchmark investors like ETFs and mutual funds have benchmark indices to track, most of which are net (total) return indices. Passive ETFs and mutual funds aim to minimize the tracking errors against their benchmark indices. Even for active funds, closet indexing is common (Cremers and Petajisto 2009, Cremers et al. 2016). As the equity index's pro-rata dividend reinvestment implies dividend repatriation, benchmark investors have particular incentives to repatriate dividends as well.

Despite equity indices prescribing dividends to be reinvested on the ex-date,[26] investors do not receive dividends until the payment date. Between the ex-date and the payment date, dividends are accrued to investors' accounts.[27] Accrued dividends are not reinvested and are in local currency. Therefore, compared with equity index treatment, accrued dividends will lead to tracking errors due to cash drag and FX fluctuations between the ex-date and the payment date. If a fund manager chooses to reinvest dividends in exactly the same way

25. Weiner (2023) Chapter 3 provides an example showing how the dividend affects the shares count in index close file before the ex-date and index open file on the ex-date.

26. Index methodologies prefer to assume all dividends are reinvested on the ex-date rather than incur the complications of allowing a time lag before reinvesting the declared dividends on the payment date. See FTSE (2023) Section 4.5.1.

27. Dividend accrual is reflected in the fund NAV calculation and recorded under the receivables in the financial statement.

as the underlying index methodology on the ex-date, he will need to borrow money, which incurs additional funding costs. Alternatively, he can wait until dividends are paid and then act. Depending on the institutional setup, the dividends may appear on the fund's available cash account on or shortly after the payment date.[28] Once the cash hits the account, the fund manager has incentives to act fast, as further delay may lead to further tracking errors. Regarding the exact implementation of reinvestment, the fund manager can repatriate to other currencies and reinvest directly into the underlying stocks. Or, more commonly, he can repatriate the dividends back to the fund currency and use futures to establish effective exposures. Doing so is more cost-effective. In either case, a proportion of the dividends are predictably repatriated from the currency that pays the dividends and converted into other currencies.

I define the *dividend repatriation channel* as investors' predictable repatriation of dividends from the dividend currency to other currencies shortly after receiving the dividend payments. This channel differs from month-end or quarter-end rebalancing, as the timing is different, i.e., the dividend repatriation channel is in the near term. This also differs from the portfolio rebalancing due to risk-averse investors' portfolio optimization as in Camanho et al. (2022). In my paper, the dividend repatriation channel is due to benchmark investors' minimization of tracking errors against global equity indices. Such dividend repatriation is mechanical and hence informationless.

Figure 1.6 uses detailed daily positions from a specific ETF, First Trust Developed Markets ex-US AlphaDEX® Fund (FDT), to illustrate the dividend repatriation channel. Launched in April 2011 and issued by First Trust, FDT is a passive global equity ETF tracking NASDAQ AlphaDEX Developed Markets Ex-US Index. As of December 2022, its assets under management (AUM) are 419 million USD. It has relatively clean daily cash reporting in

28. Hartzmark and Solomon (2022) notices that cash may appear on investors' accounts even after the payment date due to institutional reasons.

ETF Global and does not have frequent fund inflows/outflows (i.e., creation/redemption).[29] This case study provides a clear illustration of how the dividend repatriation channel works. Consider the period from November 30 to December 9, 2022. During this period, there are no fund inflows or outflows, no change in underlying stock positions, and no distributions to the ETF investors. Imputed from FDT's portfolio holdings and the stocks' dividend information, the fund should receive dividends from its portfolio holdings of Japanese companies from November 30 (Wednesday) to December 2 (Friday) in JPY, with the dividend payment on December 1 (Thursday) being the largest. In the meantime, dividends received in other currencies are negligible. The JPY dividends appeared in its JPY cash account on December 5 (Monday), after which the JPY cash position decreased while the USD cash position increased by a similar amount. Note that the USDJPY spot transaction follows T+2 settlement rule, i.e., the cash is delivered at time T+2 for a spot transaction done at time T. Therefore, the "sell JPY/buy USD" trade should be conducted on December 5 for the JPY cash position to decline on December 7. Such trade affects the foreign exchange rate on December 5 (Monday), which is two business days after the large dividend payment on December 1 (Thursday). This time lag is consistent with the empirical results identified in the baseline regression Eq (1.2), i.e., the point estimates are statistically significant since $t = 2$ in Table 1.6.

1.6.3 Financial Intermediaries with Limited Risk-Bearing Capacity

Accommodating benchmark investors' currency demand are financial intermediaries. Gabaix and Maggiori (2015) and Itskhoki and Mukhin (2021) highlight the central role of financial intermediaries in FX determination. He et al. (2017) and Reitz and Umlandt (2021) provides empirical evidence that financial intermediaries price FX. Importantly, financial intermediaries have limited risk-bearing capacity. Limited risk-bearing capacity may result from

29. ETF creation/redemption can either be in-kind, in-cash, or mixed, i.e., it may contain cash components in the basket. See Koont et al. (2023).

regulations (Du et al. 2018), risk management (Fang and Liu 2021), or margin constraints (Garleanu and Pedersen 2011). Sandulescu et al. (2021) shows financial intermediaries' risk-bearing capacity explains the time variation of international SDFs. As the risk-bearing capacity is limited and the balance sheet is constrained, for financial intermediaries to accommodate the currency demand, they require compensation to take the other side of the market,

Financial intermediaries are heterogeneous. They have different sophistication and different beliefs. They trade heavily among themselves. According to the latest BIS Triennial Central Bank Survey (BIS 2022), 46% of global turnover of FX are among reporting dealers[30], and 22% are between reporting dealers with non-reporting banks[31]. Broadening the definition of financial intermediaries to include arbitrage capital like hedge funds and proprietary desks, 7% of global FX turnover is between reporting dealers and hedge funds & proprietary trading firms.[32]

Unlike unexpected capital flows, in principle, dividend flows can be estimated ex-ante. This is because aggregate dividend payments are predetermined (Section 1.3) and a certain proportion of dividends are predictably repatriated shortly after dividend payment dates

30. According to BIS (2022), reporting dealers are defined as financial institutions that participate as reporters in the Triennial Survey. These are mainly large commercial and investment banks and securities houses that (i) participate in the inter-dealer market and/or (ii) have an active business with large customers, such as large corporate firms, governments and non-reporting financial institutions; in other words, reporting dealers are institutions that actively buy and sell currency and OTC derivatives both for their own account and/or to meet customer demand.

31. According to BIS (2022), non-reporting banks are smaller or regional commercial banks, publicly owned banks, securities firms or investment banks not directly participating as reporting dealers

32. According to BIS (2022), hedge funds & proprietary trading firms are (i) Investment funds and various types of money managers, including commodity trading advisers (CTAs), which share (a combination of) the following characteristics: they often follow a relatively broad range of investment strategies that are not subject to borrowing and leverage restrictions, with many of them using high levels of leverage; they often have a different regulatory mandate than "institutional investors" and typically cater to sophisticated investors such as high-net-worth individuals or institutions; and they often hold long and short positions in various markets, asset classes and instruments, with frequent use of derivatives for speculative purposes. (ii) Proprietary trading firms that invest, hedge, or speculate for their own account. This category may include specialized high-frequency trading (HFT) firms that employ high-speed algorithmic trading strategies characterized by numerous frequent trades and very short holding periods.

by benchmark investors (Section 1.6.2). Nevertheless, financial intermediaries may differ in their sophistication in collecting and processing this information. Therefore, they may have different beliefs on the FX implications. The model in Section 1.6.4 shows that heterogeneous intermediaries with limited risk-bearing capacity that meet the dividend repatriation flows from benchmark investors are the underlying reason why predetermined dividend flows move the exchange rate shortly after the payment dates.

1.6.4 Model

In this section, I present a partial equilibrium model of the currency market. This model explains the dynamics of exchange rates around dividend payment dates and dividend announcement dates. Additionally, the model has further implications for the time variation of the price impact of dividend payments.

Figure 1.7 summarizes the model ingredients graphically. There are two countries, the US and the UK. Denote the exchange rate E_t as units of USD per GBP, i.e., the strength of GBP. A negative change in E_t means GBP depreciates.

There are three periods $t = 0, 1, 2$. At time 0, UK companies announce the dividend payment in GBP, with the ex-date and the payment date both at time $t = 1$.[33] Time $t = 2$ is the long-run equilibrium, where the exchange rate is expected to revert back to the steady state $\bar{E}$ on average, i.e.,

$$\mathbb{E}_1[E_2] = \bar{E}, \quad Var_1[E_2] = \sigma_E^2 \tag{1.3}$$

Trading takes place over the time interval $[0, 1]$ at equally spaced time points $t_n = n\Delta, n = 0, ..., N$ where $N\Delta = 1$.

There are four agents: a benchmark investor following the global equity index, a noise trader, and two types of financial intermediaries with limited risk-bearing capacity in the

33. For simplicity, I combine the ex-date and the payment date together. In some countries like Switzerland, the ex-date and the payment date are only a few days apart.

currency market.

The benchmark investor mechanically follows the equity index in order to minimize the tracking errors. The global equity index methodology prescribes the reinvestment of dividends into the entire portfolio pro rata, including both in the UK and the US. Upon the dividend payment from the UK in GBP at time $t = 1$, the benchmark investor repatriates a certain proportion out of the currency, i.e., dividend flow. To do so, the benchmark investor needs to sell f GBP and buy USD at time $t = 1$, where f is a constant known at time $t = 0$. The benchmark investor does not trade before the payment date as its benchmark equity index does not change.

The noise trader has a stochastic demand for currency at time t_n, independent of everything. It buys η_{t_n} GBP and sells the equivalent amount in USD, where $\eta_{t_n} \sim N(0, \sigma_\eta^2)$. η_{t_n} can be either positive or negative. If $\eta_{t_n} < 0$, it means the noise trader sells $|\eta_{t_n}|$ GBP and buys the equivalent amount of USD. For simplicity of notation, assume that the strength of the noise trader's currency demand is such that $Var_{t_{n-1}}[E_{t_n}] = \sigma_E^2$, i.e., the volatility of exchange rates is constant over time. For this condition to hold, we need the parameter assumption $\sigma_\eta = 1/(\gamma\sigma_E)$.

The financial intermediaries are heterogeneous, with $\lambda \in [0, 1)$ proportion being sophisticated type A (e.g., hedge funds), $1 - \lambda$ proportion being unsophisticated type B (e.g., dealers).[34] Both type A and type B intermediaries are mean-variance investors with risk aversion γ. They maximize the following utility function to determine their demand for

34. This modeling device is similar to Hau (2011), in which type A intermediaries are labeled as informed arbitrageurs and type B intermediaries are labeled as uninformed liquidity providers.

GBP at time t:[35][36]

$$\max_x \mathbb{E}_t^i[(E_{t+1} - E_t)x] - \frac{\gamma}{2} Var_t[(E_{t+1} - E_t)x] = \mathbb{E}_t^i[(E_{t+1} - E_t)]x - \frac{\gamma\sigma_E^2}{2} x^2$$

This gives the following demand curve for GBP for the type i intermediary:

$$q_t^i = \frac{1}{\gamma\sigma_E^2} \mathbb{E}_t^i[E_{t+1} - E_t]$$

i.e., they trade off the expected return with the volatility, the latter of which can be interpreted as the holding cost of the position for the intermediaries.

The two types of intermediaries differ in their beliefs of expectations of the future exchange rate. Type A intermediaries have rational expectations, in the sense that their expectation of the future exchange rate is correct:

$$\mathbb{E}_t^A[E_{t+1}] = \mathbb{E}_t[E_{t+1}] \tag{1.4}$$

In particular, type A intermediaries are attentive to the dividend payments forthcoming at $t = 1$ and the associated dividend repatriation when they form their expectation of the next-period exchange rate $\mathbb{E}_t[E_{t+1}]$. Aggregating λ measure of type A intermediaries, their demand curve for currency depends on the exchange rate today and tomorrow, as in Gabaix and Maggiori (2015), Itskhoki and Mukhin (2021):

$$Q_t^A = \lambda q_t^A = \frac{\lambda}{\Gamma} \mathbb{E}_t(E_{t+1} - E_t) \tag{1.5}$$

where $\Gamma = \gamma\sigma_E^2$ represents the (inverse) risk-bearing capacity of the financial intermediary

35. For simplicity, I assume gross interest rates in both countries are equal to 1. In this model, currencies are synonyms for bonds.

36. Here, the subscript $t + 1$ means the next period following time t, i.e., the next period of t_n is t_{n+1}, the next period of $t = 1$ is $t = 2$.

sector, with smaller Γ being the larger risk-bearing capacity. Type A intermediaries will demand more GBP if they expect GBP to appreciate against USD, which makes buying GBP and selling USD a profitable trade. On the other hand, if they expect GBP to depreciate in the future due to the benchmark investor's selling at $t = 1$, they will sell GBP beforehand.

In contrast, type B intermediaries are less sophisticated. They do not understand the implications of dividend payments on exchange rates. Therefore, type B intermediaries' expectation of the next period's exchange rate is always the long-run equilibrium exchange rate,

$$\mathbb{E}_t^B[E_{t+1}] = \bar{E} \tag{1.6}$$

Aggregating $1 - \lambda$ measure of type B intermediaries, their currency demand depends on the deviation of the exchange rate at time t against the long-run equilibrium exchange rate, as in Camanho et al. (2022):

$$Q_t^B = (1 - \lambda)q_t^B = \frac{1 - \lambda}{\Gamma}(\bar{E} - E_t) \tag{1.7}$$

Given the long-run equilibrium exchange rate $\bar{E}$, type B intermediaries' demand only depends on the exchange rate today. If the current exchange rate is lower than $\bar{E}$, type B intermediaries will buy GBP and sell USD.

The following proposition summarizes the equilibrium exchange rate dynamics and the intermediaries' positions of GBP:

Proposition 1. *In equilibrium, the exchange rate at time $t_n = n\Delta, n = 0, ..., N$ is*

$$E_{t_n} = \bar{E} - \lambda^{N-n}\Gamma f + \Gamma\eta_{t_n}; \quad \mathbb{E}[E_{t_n}] = \bar{E} - \lambda^{N-n}\Gamma f \tag{1.8}$$

Before the payment date at time $\{t_n\}_{n=0}^{N-1}$, type A intermediaries gradually build up the short *positions:*

$$Q^A_{t_n} = \lambda\left(-\lambda^{N-n-1}(1-\lambda)f - \eta_{t_n}\right); \quad \mathbb{E}[Q^A_{t_n}] = -\lambda^{N-n}(1-\lambda)f \tag{1.9}$$

while type B intermediaries take long positions:

$$Q^B_{t_n} = (1-\lambda)\left(\lambda^{N-n}f - \eta_{t_n}\right); \quad \mathbb{E}[Q^B_{t_n}] = \lambda^{N-n}(1-\lambda)f \tag{1.10}$$

At the payment date $t_N = 1$, the benchmark investor sells GBP while both intermediaries buy:

$$Q^A_1 = \lambda(f - \eta_1), Q^B_1 = (1-\lambda)(f-\eta_1); \quad \mathbb{E}[Q^A_1] = \lambda f, \mathbb{E}[Q^B_1] = (1-\lambda)f \tag{1.11}$$

Figure 1.8 plots the expected value of the exchange rate and positions of type A and type B intermediaries in equilibrium. Expecting GBP to depreciate as in Eq (1.8), type A intermediaries gradually build up short positions in GBP by selling GBP to type B intermediaries before the dividend payment date. Type B intermediaries are willing to buy GBP because their belief is different. They expect the next period exchange rate will revert back to the steady state $\bar{E}$, as specified in Eq (1.6). According to Eq (1.9), the size of short positions for intermediaries is larger when it is closer to the payment date. With the announcement date being far away from the payment date (Table 1.2), the size of short positions by intermediaries around the dividend announcement date is negligibly small. In contrast, the largest short position is taken by the type A intermediaries immediately before the dividend payment date. This is because deploying capital to take positions is costly, as reflected by the negative variance part in the mean-variance utility function. Since Type A intermediaries do not aggressively take short positions far in advance, we should observe a limited anticipation effect before the payment date. This is precisely what I show empirically in Section 1.5.

In the exchange rate dynamics related to dividends, the payment date effect, the anticipation effect, and the announcement date effect are of particular interest. The following proposition shows the magnitude of these two effects:

Proposition 2. *Define the payment date effect of dividend flows on the foreign exchange rate as the expected FX change upon the dividend payment date, i.e.,* $\mathbb{E}[E_1 - E_{t_{N-1}}]$. *Define the anticipation effect as the expected cumulative FX change from the announcement date to the date prior to payment date, i.e.,* $\mathbb{E}[E_{t_{N-1}} - \bar{E}]$. *Define the announcement date effect as the expected FX change upon the dividend announcement date, i.e.,* $\mathbb{E}[E_0 - \bar{E}]$.

1. *Payment date effect:*
$$\mathbb{E}[E_1 - E_{t_{N-1}}] = -(1-\lambda)\Gamma f \tag{1.12}$$

2. *Anticipation effect:*
$$\mathbb{E}[E_{t_{N-1}} - \bar{E}] = -\lambda\Gamma f \tag{1.13}$$

3. *Announcement date effect:*
$$\mathbb{E}[E_0 - \bar{E}] = -\lambda^N \Gamma f \tag{1.14}$$

As Proposition 2 shows, dividend flow moves the foreign exchange rate at the payment date. The magnitude of the payment date effect is increasing in the proportion of type B intermediaries $(1-\lambda)$. In other words, the less the arbitrage capital (aka. λ proportion of type A), the more pronounced is the payment date effect. If every intermediary is a forward-looking arbitrageur, i.e., $\lambda = 1$, there are no counterparties for them to trade with before the payment date on average, as the demand from the noise trader is 0 in expectation. The exchange rate before the payment date will immediately adjust for there to be no trade. In this case, the payment date effect will be zero. Therefore, to have a significant dividend payment date effect, we need type B intermediaries, which do not fully understand the implications of dividend payments on FX. When $\lambda = 0$, i.e., all intermediaries are type B, the model is equivalent to the model in which the capital flow f is unexpected. In this case, the price impact of capital flow is the largest at $-\Gamma f$.

In contrast, the announcement date effect increases with the proportion of type A intermediaries λ: the greater the amount of forward-looking arbitrage capital, the more pronounced the announcement date effect. In addition, the announcement date effect decreases with the time gap between the announcement date and the payment date N. Empirically, I show in Sections 1.5.1 and 1.5.2 that the payment date effect is economically large and statistically significant, while the announcement date effect is small and insignificant.

With a reasonably small λ and a large N, the model quantitatively explains the significant payment date effect, the limited anticipation effect, and the negligible announcement date effect, as illustrated in Figure 1.8. Here, for the calibration of λ, I use the point estimates in the baseline regression in Table 1.6 Panel *Two-Way Fixed Effects with Controls*. For the payment date effect, I take the point estimate with the largest magnitude at $t = 8$, which is 6.48 basis points. This is to account for the potential delayed response. As the anticipation effect estimates are insignificant, I take the average of point estimates from $t = -10$ to $t = -2$ to increase precision. This gives 2.75 basis points. Therefore, $\lambda = 2.75/(2.75 + 6.48) \approx 0.30$.

If dividends are recurring events, why don't type B intermediaries learn from the FX dynamics and correct their beliefs? Firstly, inferring from the FX dynamics jointly with dividend payments requires expertise, which varies significantly among financial institutions. Secondly, financial intermediaries may have different objectives. In the model, in spite of the short-term loss from long positions in GBP before the payment date, type B intermediaries eventually profit from these positions as the exchange rate typically reverts to the steady state in the long run. Thirdly, since many other factors affect exchange rates, the signal-to-noise ratio of dividend payments is relatively low. The identified magnitude of the payment date effect of large dividend payments in Section 1.5 is around 5 to 10 basis points. In contrast, the daily volatility of G10 currencies (against USD) is 68 basis points in the sample period from January 2001 to June 2023. Therefore, learning this effect and correcting their priors may take a long time.

The model has further implications for the time variation of dividend flows' price impact on the foreign exchange rate. If the intermediary mix λ is relatively stable, the price impact of dividend flows depends on the time variation of (inverse) risk-bearing capacity parameter $\Gamma := \gamma\sigma_E^2$. The risk aversion γ can be interpreted as the balance sheet constraints of financial intermediaries, while σ_E^2 stands for the FX market volatility. When the balance sheet constraints are tight, or the market volatility is high, the risk-bearing capacity of intermediaries will be low. I use the intermediary capital ratio and the CIP deviations to proxy for the balance sheet constraints, as the intermediary capital ratio is the cause while the CIP deviations are the result. I use the currency implied volatility to proxy for the FX market volatility, as it is forward-looking. I summarize the implications of time variation of dividend impact on the foreign exchange rate in the following proposition:

Proposition 3. *The price impact of dividend flows on the foreign exchange rate is larger, if*

1. *the intermediary capital ratio is lower*
2. *the CIP deviations are larger*
3. *the currency implied volatilities are higher*

1.7 Time-Variation in the Price Impact of Dividend Flows

In this section, I empirically test three implications of the limited risk-bearing capacity of financial intermediaries. I find that the price impact of dividend flows is larger when the intermediaries' risk-bearing capacity is lower, e.g., when the intermediary capital ratio is lower, the covered interest parity (CIP) deviations are larger, and the currency implied volatilities are higher.

Consistent with the pattern established in Section 1.5.1, I focus on the two-day cumulative change after dividend payments in this section. The short-run effect is closer to the essence of dividend repatriation as highlighted in Section 1.6.2. Other horizons give similar results, though the power of the test may decrease as the horizon increases. The main specification is similar to Eq (1.2), as follows:

$$\Delta_2 e_{k,t+2}^{US/LC} := \ln E_{k,t+2}^{US/LC} - \ln E_{k,t-1}^{US/LC} = \alpha + \beta DivOut_{k,t} + Controls + \gamma_k + \xi_t + \epsilon_{k,t+2} \quad (1.15)$$

The key variable $DivOut_{k,t}$ is country i's dividends paid out to foreign investors on date t normalized by the previous year-end local stock market capitalization. Both the numerator and denominator are in the local currency. Therefore, there is no foreign exchange rate involved in the construct of $DivOut_{k,t}$. Dividends paid out to foreign investors are calculated using total dividend payments from Compustat Global/CRSP, multiplied by the foreign ownership calculated in Appendix 1.10.2. As Figure 1.9 shows, this calculation matches the dividends imputed from the Balance of Payments closely. Scaling by the foreign ownership is to control for its increasing trend, as higher foreign ownership implies potentially larger dividend repatriation flows f, given the same amount of dividend payments. Normalization by the previous year-end local stock market capitalization makes $DivOut_{k,t}$ stationary, as both dividends and stock market capitalization have grown significantly over the past 20 years.

1.7.1 Intermediary Capital Ratio

The intermediary capital ratio can be used as a proxy for the balance sheet constraint of financial intermediaries. As in He et al. (2017), I define intermediary capital ratio as the New York Fed's primary dealers' market equity divided by market equity plus their aggregate book debt. The New York Fed's primary dealers are the New York Fed's trading counterparties in implementing monetary policy. The primary dealers are large financial institutions[37], many of which are active in the G10 currency market. Therefore, their capital ratio should be relevant for the G10 currency market. Reitz and Umlandt (2021) refines the intermediary capital ratio for the currency markets using the balance sheet data of the top three foreign exchange dealers. Their measure is highly correlated with He et al. (2017), with the correlation being 0.90 from 1999 to 2017, when Reitz and Umlandt (2021) sample ends. The results in this section are qualitatively and quantitatively similar if using Reitz and Umlandt (2021)'s measure.

Table 1.7 Panel B confirms Proposition 3.1. Column 1 reiterates the findings in Section 1.5 using continuous variable $DivOut_{k,t}$ in Eq (1.15). The price impact coefficient implies 1% local stock market capitalization paid out to foreign investors as dividends will lead to the local currency depreciation against USD by 0.806% in two days time after the payment date. Column 2 and Column 3 are split sample regressions. Column 2 is over the subsample where the intermediary capital ratio is greater than the median. This is when the balance sheet constraint is looser. The estimated price impact coefficient is -0.192 and statistically insignificant. Column 3 is over the subsample where the intermediary capital ratio is smaller than the median. This is when the balance sheet constraint is tighter. The estimated price impact coefficient is -1.209 and statistically significant. Column 4 adds the interaction term between $DivOut_{k,t}$ and the subsample dummy variable in addition to first-order terms, with fully saturated fixed effects. It shows that the difference in the price impact coefficient in Column 2 and Column 3 is economically large and statistically significant, i.e., when the capital ratio is lower, the price impact of dividend flows on the foreign exchange rate is larger.

37. As of 2023, the primary dealers include ASL Capital Markets, Bank of Montreal, Bank of Nova Scotia, BNP Paribas Securities, Barclays Capital, BofA Securities, Cantor Fitzgerald & Co., Citigroup Global Markets, Daiwa Capital Markets America, Deutsche Bank Securities, Goldman Sachs & Co., HSBC Securities (USA), Jefferies, J.P. Morgan Securities, Mizuho Securities USA, Morgan Stanley & Co., NatWest Markets Securities, Nomura Securities International, RBC Capital Markets, Santander US Capital Markets, Societe Generale, TD Securities (USA), UBS Securities, Wells Fargo Securities

1.7.2 Deviations from Covered Interest Rate Parity

Another proxy of the balance sheet constraints of financial intermediaries is the deviations from covered interest rate parity (CIP). Traditionally, CIP is a textbook example of no-arbitrage condition. It requires the US dollar interest rate in the cash market to be the same as the synthetic dollar interest rate, which borrows in foreign currency and use FX swap to transform into USD. Since the 2007-2008 Global Financial Crisis (GFC), the CIP deviation has been persistent. Duffie (2017), Du et al. (2018) find this is the result of the post-GFC regulatory reforms in the banking sector, especially the non-risk-weighted capital requirements in the form of the leverage ratio or supplementary leverage ratio. Following the GFC, new regulations (e.g., the Basel III leverage ratio rule and the U.S. supplementary leverage ratio) were introduced that require banks to maintain a minimum capital ratio against all assets, regardless of their risk characteristics. This limits global banks' capacity to arbitrage. Du et al. (2023) shows that CIP deviations are correlated with the other types of near-arbitrages, including bond-CDS basis, the CDS-CDX basis, the USD Libor tenor basis, 30-year swap spreads, the Refco-Treasury spread, the KfW-Bund spread, and the asset-swapped TIPS/Treasury spread. Therefore, I use CIP deviation as a barometer for the intermediaries' balance sheet constraints, or more broadly, the scarcity of arbitrage capital.

Following the literature, I measure the CIP deviation using the cross-currency basis against USD, i.e.,

$$x_t^k = i_t^{US} - (i_t^k - \rho_t^k)$$

where i_t^{US} is the US dollar interest rate in the cash market, $(i_t^k - \rho_t^k)$ is the synthetic US dollar interest from the FX swap market. $\rho_t^k = (s_t^k - f_t^k)^4$ is the annualized forward premium, where s_t^k is the log spot exchange rate and f_t^k is the log 3-month forward outright, both in terms of units of USD per local currency.

Table 1.7 Panel B confirms Proposition 3.2. Column 1 is the full sample results. Column 2 and Column 3 are split sample regressions. On the subsample where the absolute value of the

CIP deviation is lower than the median within currency, the price impact coefficient is -0.302 and statistically insignificant. This is when the balance sheet constraints are less binding. On the subsample where the absolute value of the CIP deviation is higher than the median within currency, the price impact coefficient is -1.259 and statistically significant. This is when the balance sheet constraints are more stringent. Adding the interaction term between $DivOut_{k,t}$ and the subsample dummy variable in addition to first-order terms, Column 4 confirms the difference in price impact coefficient in Column 2 and Column 3 is not only economically large but also statistically significant. That is to say, when the balance sheet constraints are more stringent, the price impact of dividend flows on the foreign exchange rate is larger.

1.7.3 Currency Implied Volatility

In addition to the risk aversion coefficient γ, the FX volatility σ_E also affects the intermediary risk-bearing capacity Γ. In reality, this can stem from financial intermediaries' risk management practice in the form of value-at-risk (VaR) constraints (e.g., Fang and Liu (2021)). VaR constraints are widely used in the financial industry, including banks, hedge funds, etc. As higher volatility translates into tighter VaR constraints, the intermediaries' risk-bearing capacity is lower.

The FX volatility in the model in Section 1.6.4 is next-period volatility. Therefore, to proxy σ_E, I use the FX implied volatilities which is forward-looking. I use 6-month tenor as it strikes a balance between short-term and long-term volatility. Using other tenors or realized volatility gives similar results.

Table 1.7 Panel C confirms Proposition 3.3. Column 1 is the full sample results, while Column 2 and Column 3 are the results for split sample regressions. When the implied volatility is lower than the median within currency, the price impact coefficient is -0.359 (Column 2). This is when the intermediary risk-bearing capacity is larger. When the implied volatility is lower than the median within currency, the price impact coefficient is -1.290 (Column 3). This is when the intermediary risk-bearing capacity is smaller. Adding the interaction term between $DivOut_{k,t}$ and the subsample dummy variable in addition to first-order terms, Column 4 confirms the difference in price impact coefficient in Column 2 and Column 3, -0.931, is economically large but also statistically significant. Therefore, at a time when the currency implied volatility is higher, the price impact of dividend flows on the foreign exchange rate is larger.

1.8 Implications for International Finance

In this section, I discuss the implications of my paper. First, I provide a back-of-the-envelope calculation of the price multiplier in the FX market, compare it with other estimates in the literature, and link it to the inelastic market hypothesis developed by Gabaix and Koijen (2021). Second, I discuss how the price impact estimates are useful to shed light on intermediaries' capital requirements. Third, I present evidence that the price impact of dividend flows is larger in the free-floating FX regime than other regimes.

1.8.1 FX Elasticity

The price impact coefficient estimated using Eq (1.15) implies 1% of local stock market capitalization paid out to foreign investors in local currency as dividends will lead to the local currency depreciation against USD by 0.806% in two days time after the payment date (Table 1.7 Panel A Column 1). At the end of 2022, the average stock market capitalization in non-US G10 countries is 2,681 billion USD. Expressed in semi-multiplier,[38] this implies $33(= 1\%/0.806 \times 2681)$ billion USD-equivalent dividends payments to foreign investors are associated with 1% G10 currency movement against USD within two days.

38. Semi-multiplier is defined as $d \ln E/dQ$, where E is the foreign exchange rate against USD and the capital flow Q is expressed in USD-equivalent amount.

However, not all dividends paid out to foreign investors in the local currency are repatriated immediately. To have a sense of the magnitude of actual dividend repatriation flows, we need to estimate the *dividend repatriation intensity* in the short run. In Section 1.6.2, I argue the short-run effect of dividend payments on the foreign exchange rate is most likely due to benchmark investors' dividend repatriation channel. Ideally, one should collect ETF and mutual fund holding data globally. Due to data limitations, I use US-domiciles ETFs and mutual funds in the following back-of-envelop calculation. This should provide a reasonable estimate, as the total assets under management (AUM) of US-domiciled ETFs and mutual funds far outsize those in other countries. Using Morningstar data (Figure 1.1), as of 2020 year-end, US-domiciled ETFs hold 3.2% of the local stock market capitalization, average across non-US G10 countries. In addition, US-domiciled mutual funds hold 6.6% of the local stock market capitalization. In total, US-domiciled benchmark investors hold 9.8% of local stock market capitalization. Using the data in Appendix 1.10.2, the foreign ownership across non-US G10 countries is 40.3% as of 2020.

Therefore, $\approx 24.3\%(= 9.8\%/40.3\%)$ of the dividends paid to foreign investors are paid to the US-domiciled benchmark investors, who are likely to be repatriated out of the local currency.[39] With $\approx 24.3\%$ dividend repatriation intensity of foreign investors extrapolated to 2022,[40] 33 billion USD dividends paid to foreign investors is translated to 8 billion USD dividend repatriation flows out of the local currency. To conclude, on average, dividend flows of $0.30\%(= 1\%/0.806 \times 24.3\%)$ of local stock market capitalization move the G10 currency by 1%. In terms of semi-multiplier, $\$8.1(= 1\%/0.806 \times 24.3\% \times 2681)$ billion dividend flows move the G10 currency by 1% vis-à-vis USD.

Table 1.8 compares my estimates with the others in the literature. The existing papers

39. This back-of-the-envelope estimate ignores non-US based ETFs and mutual funds, though they are much smaller than US-domiciled counterparts. In addition, US-domiciled benchmark investors may keep a certain proportion of dividends reinvested in the local stock markets.

40. Here, I extrapolate from 2020 to 2022 as my sample of US-domiciled ETFs/mutual funds from Morningstar ends in 2020Q4.

often rely on ad-hoc normalization, including GDP, M2, market capitalization, etc. Therefore, I convert the numbers in these papers to the semi-multiplier, i.e., the dollar amount of flows that can move the exchange rate by 1%. Though estimates differ in types of flows and currencies, my estimates generally fall in the ballpark of the existing ones in terms of order of magnitude. For the developed market (DM) currencies, the closest estimate to mine is Camanho et al. (2022). Recently, Camanho et al. (2022) uses GIV on rebalancing flow for mutual funds domiciled in the US, the UK, Eurozone, and Canada. They estimate that \$5.3bn to \$7.1bn equity flow is associated with 1% US dollar movement.[41] Their mutual fund rebalancing flows are unexpected flows, while the dividend flows I use are predetermined. Hau et al. (2010) uses the MSCI Global Equity Index redefinition from market capitalization to freely floating in 2001 and 2002, and estimates that \$2.6bn equity flow moves the exchange rate by 1% against USD over a 6-day window around the announcement date across 33 currencies (developed market currencies & emerging market currencies).[42] Their estimate is about the announcement date effect while my estimate is about the payment date effect. In Evans and Lyons (2002) estimate that a US\$1.9 billion FX order flow moves the Deutsche mark (DEM) exchange rate against USD by 1%.[43] The order flows contain contemporaneous information about exchange rates while dividends do not.

For the emerging market (EM) currencies, Pandolfi and Williams (2019) uses the 10% cap rule in J.P. Morgan Government Bond Index–Emerging Markets Global Diversified (GBI-EM Global Diversified) that the benchmark weight of any single country cannot exceed 10% of the index at the beginning of each month, inducing monthly rebalancings for a purely mechanical reasons. Their estimate implies \$1.4bn move the local currency against USD by 1% on average across 16 EM currencies.[44] Broner et al. (2021) uses the unexpected

41. p5262-5264.

42. p1699 estimates that an (uninformative) capital flow of US\$1 billion therefore amounts to an average appreciation of 0.38% against USD.

43. p178: \$1 billion of net dollar purchases increases the Deutsche Mark price by 0.54 percent.

44. p393 Table 6 estimates that 1% inflow, relative to the market value of the sovereign bonds, leads to a

announcement of index inclusion into local-currency sovereign debt indexes of Citigroup WGBI and JP Morgan GBI-EM, and estimates \$5bn inflow leads to 1% local currency appreciation against USD in the two days following the announcement.[45] However, they find no effect during the implementation period between 2 and 6 months after the announcement date. Recently, Aldunate et al. (2022) uses Chilean pension funds flows induced by a Chilean financial advisor' uninformed market timing recommendations. Their estimate implies that \$1.4bn produces a depreciation of the Chilean peso against US dollar by 1%.

1.8.2 Capital Regulation

Regulations on global banks affect their risk-taking appetite. Even for arbitrage capital like hedge funds to size up their positions, they often need funding from banks, hence taking space in banks' balance sheets. Since the Global Financial Crisis (GFC), regulations on intermediaries' balance sheets have tightened considerably (Du et al. (2023)). This is consistent with the pattern we see in Figure 1.17 that dividends have a larger price impact on exchanges than pre-GFC.[46] As the CIP deviation can be used as a proxy for balance sheet constraints, this is also consistent with the pattern in Table 1.7 Panel B.

On the other hand, Table 1.7 Panel A shows that a higher intermediary capital ratio in terms of equity/asset ratio (He et al. (2017)) helps alleviate the price impact of dividend flows on exchange rates. To quantify how the intermediary capital ratio affects the dividend price impact coefficient, I run the following regression with the term of capital ratio interacted with dividends paid out to foreign investors, in addition to first-order terms:

$$\Delta_2 e_{k,t+2}^{US/LC} = \alpha + (\beta_0 + \beta_1 CR_t) \times DivOut_{k,t} + Controls + \gamma_k + \xi_t + \epsilon_{k,t+2} \tag{1.16}$$

close to 0.42% appreciation against the dollar in the exchange rate. I scale back this estimate by the market value of the sovereign bonds \$60.12bn in their Table 1, i.e., $(1\%/0.42\%) \times (60.12 \times 1\%) = 1.4$.

45. p17 Fig. 11 estimates 1.1% inflow, relative to GDP, leads to a 1% appreciation in the local currency against USD. I scale back this estimate by the nominal GDP in USD of the event dates.

46. The same pattern also holds if using $DivOut_{k,t}$ as RHS variable instead of $\mathbb{D}_{i,t}$.

The parameters of interest are β_0, β_1. The results are reported in Table 1.7 Panel A Column 5. The sample average capital ratio $\overline{CR}$ is 7.38%, while 1 standard deviation $std(CR)$ is 3.1%. At $\overline{CR}$, the implied price impact coefficient is $\beta = -2.123 + 20.513 \times 7.38\% = -0.609$. At $\overline{CR} - std(CR)$, the implied price impact coefficient becomes $\beta = -2.123 + 20.513 \times (7.38\% - 3.18\%) = -1.26$. That is to say, one standard deviation decrease in the intermediary capital ratio will double the price impact of flows. Consistent with my estimates, Bippus et al. (2024) also finds banking flows on exchange rates are state-dependent, with effects twice as large when banks' capital ratios are one standard deviation below average.

1.8.3 FX Regimes

How capital flows affect exchange rates may depend on the FX regimes. If a currency is in a non-free-floating regime, central banks may need to conduct foreign exchange interventions to maintain the FX regimes. In this section, I present evidence on how the price impact of dividends on the foreign exchange rate differ in different FX regimes.

Ilzetzki et al. (2019) classifies currencies into 15 fine classifications from 1940 to 2019. Relevant to G10 currencies are the following regimes: pre-announced peg (2), de facto horizontal band $\leq 2\%$ (6), de facto crawling band $\leq 2\%$ (8), moving band $\leq 2\%$ (11), managed floating (12) and freely floating (13). I extend the last observation of classification to date. As Figure 1.10 shows, over the sample period since 2001, AUD, EUR, JPY, and USD have always been in the freely floating regime, NZD has always been managed floating (anchoring to AUD), and NOK has always been de facto moving band $\pm 2\%$ against Euro. CAD switched from de facto moving band ($\pm 2\%$ band against US dollar) to freely floating in June 2002. GBP switched from de facto moving band ($\pm 2\%$ band against Euro) to freely floating in January 2009. SEK switched from de facto horizontal band ($\pm 2\%$ band against Euro) to de facto moving band ($\pm 2\%$ band against Euro) in September 2008. CHF switched to pegging to Euro during September 2011 to January 2015, while in other time, de facto moving band ($\pm 2\%$ band against Euro). In the sample, the number of observations in freely floating regime are similar to the number of observations in other regimes. Therefore, I estimate Eq(1.15) for non-freely-floating regimes vs freely floating regime.

Table 1.9 Column 1 is the full sample results. Column 2 and Column 3 are split sample regressions. On the subsample of non-freely-floating regimes, the price impact coefficient is -0.353 and statistically insignificant. On the subsample of the freely-floating regime, the price impact coefficient is -1.689 and statistically significant. Adding the interaction term between $DivOut_{k,t}$ and the subsample dummy variable, Column 4 confirms the difference in price impact coefficient in Column 2 and Column 3 is not only economically large but also statistically significant. That is to say, the price impact of dividend flows on the exchange rate is larger in the freely floating regime than in other FX regimes. Consistent with my estimates, Beltran and He (2024) also finds that countries with a free-floating exchange rate regime (free floaters) are more than three-fold more effective at stabilizing exchange rates than are countries with a managed exchange rate regime.

1.9 Conclusion

In this paper, I show that predetermined dividends move the foreign exchange rate around the payment dates. In contrast, the anticipation effect before the payment date is limited and the announcement date effect is negligible. This empirical pattern informs us about the interaction between the benchmark investors and financial intermediaries. On the one hand, benchmark investors predictably repatriate dividends received in local currency shortly afterward. On the other hand, financial intermediaries with limited risk-bearing capacity and heterogenous beliefs give rise to FX dynamics.

Dividend payments are recurring and frequent events, compared to other one-off events like changes to indices. They can be a valuable tool in the international economists' toolbox. For example, in this paper, I use dividend flows to estimate their price impact on the foreign exchange rate at different times and under different FX regimes. As a specific type of capital flow, its predeterminedness may serve as an instrument for other capital flows.

As the FX market is often claimed to be the largest and the deepest market in the world,[47] the price effect of dividend flows and other capital flows on exchange rates appears to be very big, given the magnitude of cross-border flows like trade flows.[48] This is similar in essence to the inelastic market hypothesis, pioneered by Gabaix and Koijen (2021). In models that feature financial intermediaries' roles in FX determination, it is intermediaries' limited risk-bearing capacity that determines the elasticity of the foreign exchange rate to capital flows. That being said, reconciling the price impact estimates with other cross-border macro variables in a quantitative model is left to future research.

47. `https://www.cmegroup.com/education/courses/introduction-to-fx/what-is-fx.html`

48. It is worth noting that a significant portion of trade flows are invoiced in USD. Therefore, their FX impact may not be as big at face value.

Figure 1.1: U.S.-Domiciled ETFs and Mutual Funds:
Foreign Holdings as Share of the Local Stock Market

This figure shows the market value of US-domiciled ETFs and mutual funds equity holdings as a percentage of each country's aggregate market capitalization. The holdings of US-domiciled ETFs and mutual funds are from Morningstar with asset class being Equity or REITs. For ETFs, the sample period is from 2011 to 2020. For mutual funds, the sample period is from 2002 to 2020. The year-end aggregate market capitalization for each country is from Bloomberg.

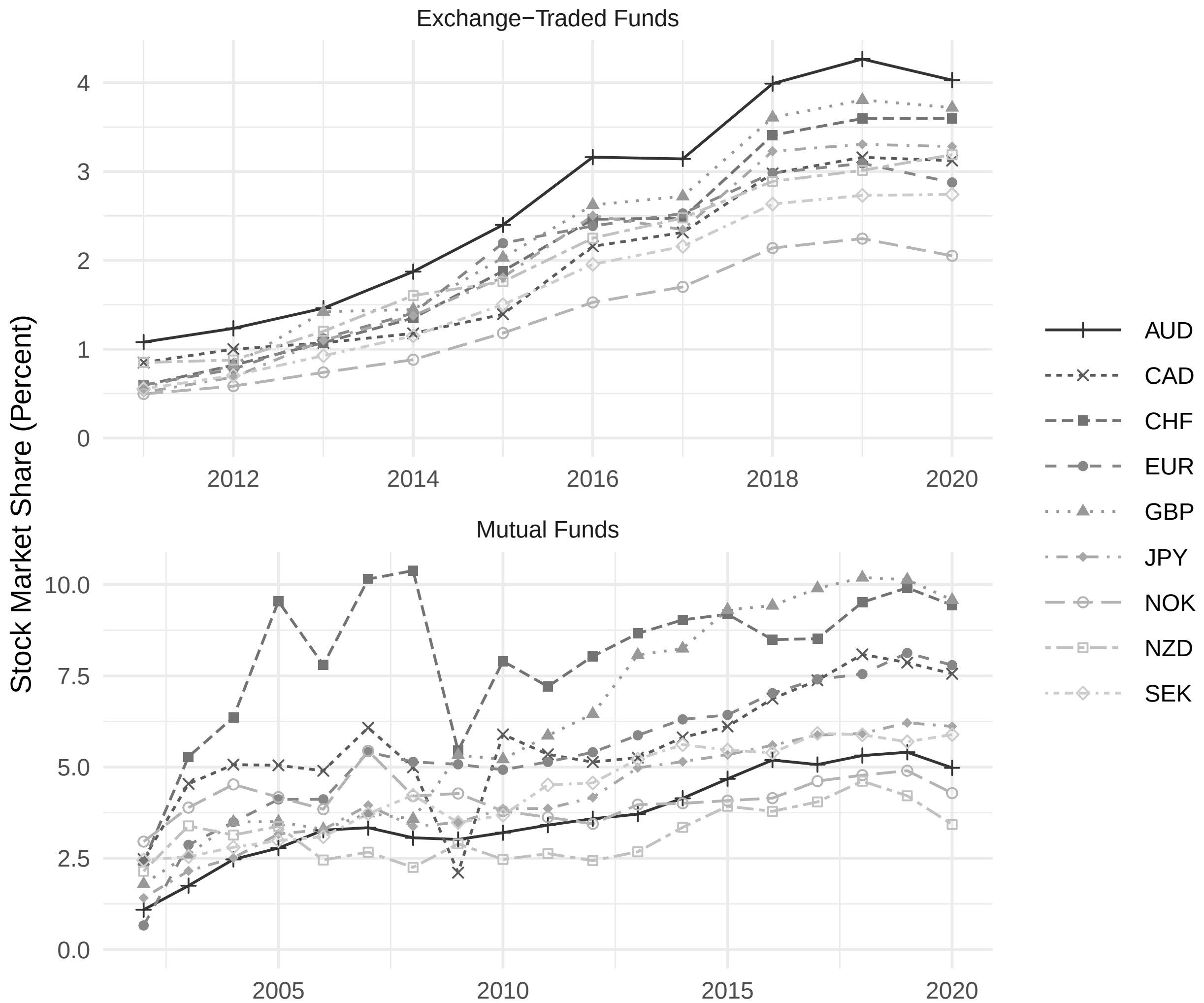


Figure 1.2: Time-Series of Cash Dividend Payments

This figure shows the dividend payments in G10 countries/currency areas from January 2018 to December 2022. I focus on cash dividends and keep common/ordinary shares that are primarily listed in a country/currency area. Dividends are aggregated to payment dates and converted to billion USD using the prevailed exchange rates on the payment dates.

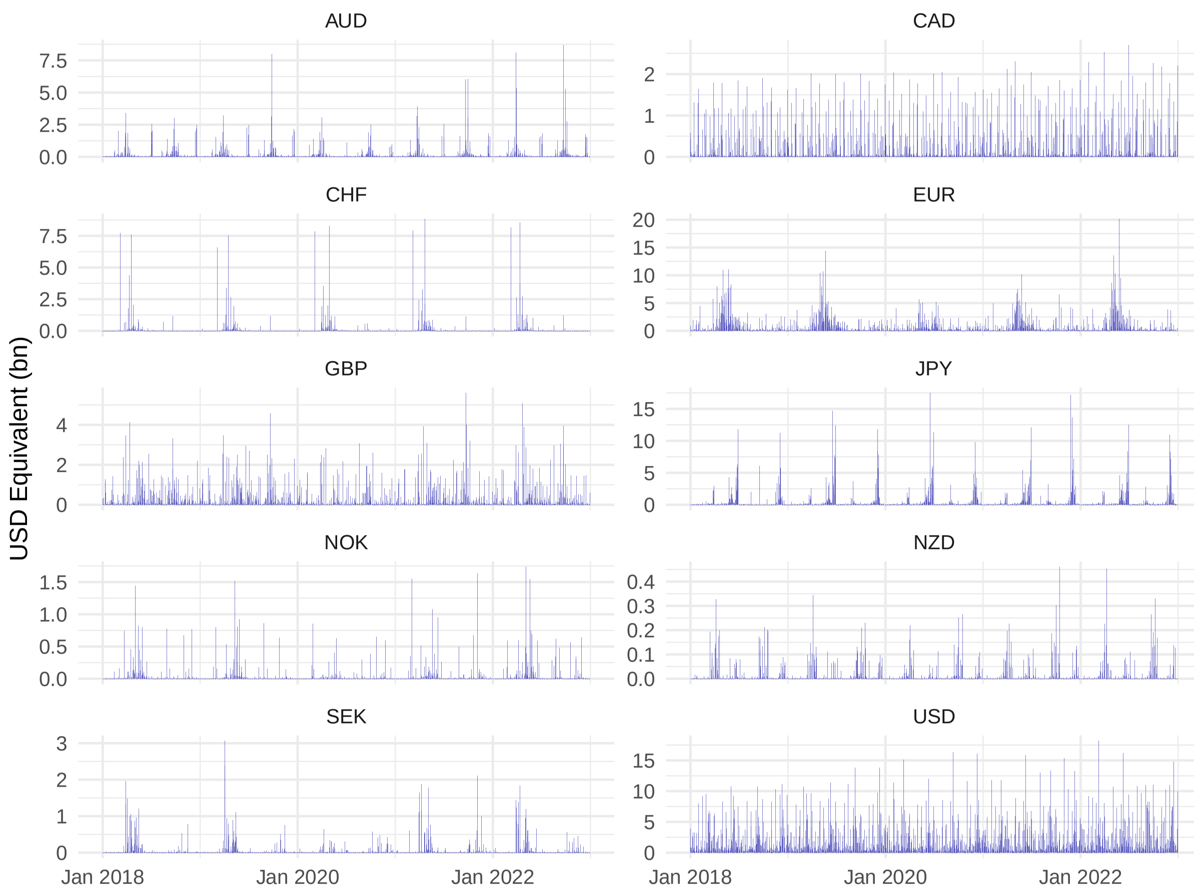


Figure 1.3: Cumulative Return on the Dividend-Based Currency Strategy

This figure shows the cumulative log returns of the dividend-based currency strategy in percentage points, both before the transaction cost (blue line) and after the transaction cost (orange line). The transaction cost, i.e., bid-ask spread, is assumed to be 1 basis point for all currencies at all times. The dividend-based currency strategy takes the following form: for each country/currency area k and date t, if in the previous l days, the combined dividend payments in the country k rank in its top p-percentile in the rolling 1-year window, then we sell currency k against USD, and hold the position for one day. If there are several currencies that satisfy this criterion, then each position is of \$1 size. The excess return on date t is calculated from summing across excess returns for each position. In this figure, $l = 2, p = 5\%$. The sample period is from January 2001 to June 2023.

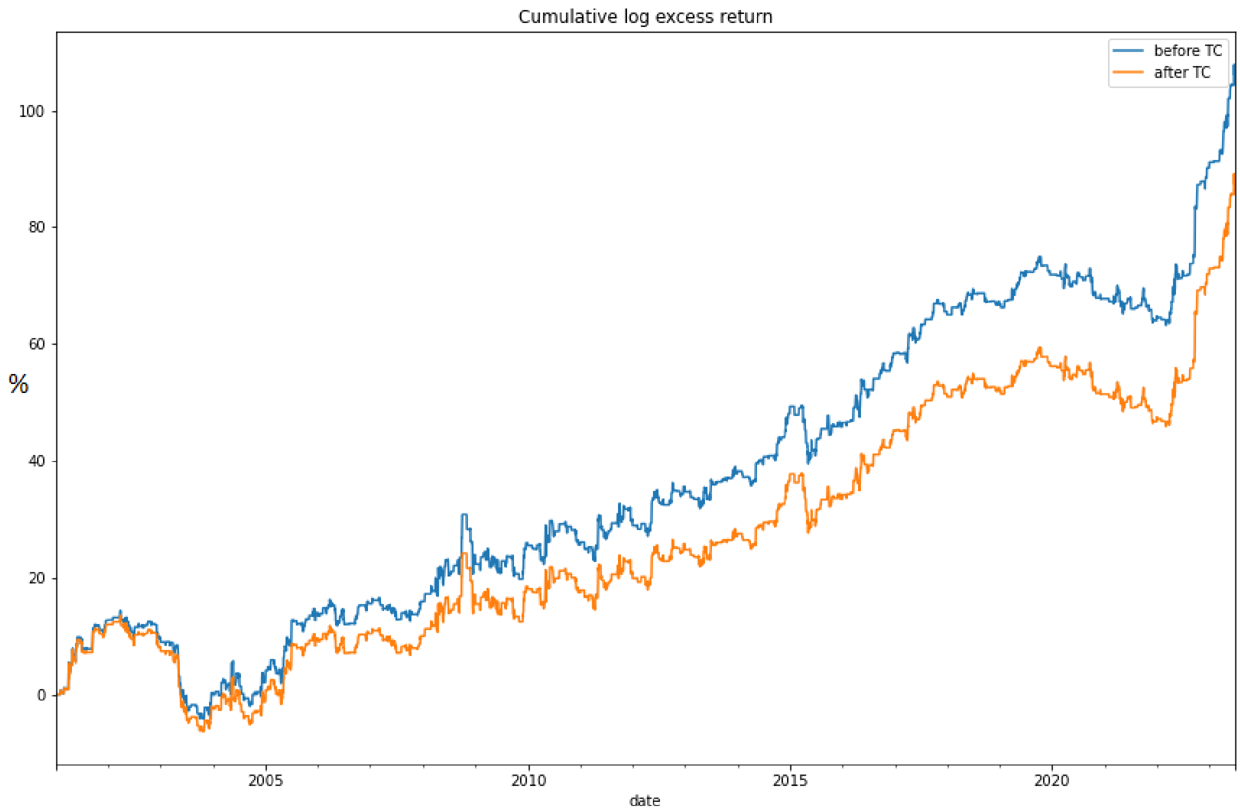


Figure 1.4: Price Impact of Large Dividends on the Foreign Exchange Rate Around Dividend Payment Dates

This figure shows the coefficients β_h estimated in the baseline regression Eq (1.2) with controls, the currency fixed effect, and the time fixed effect. Dividends are aggregated from the company level to the currency level by the payment dates. The controls include stock market returns and FX implied volatilities. The sample period is from January 2001 to June 2023. The standard errors are two-way clustered at the date level and the currency level.

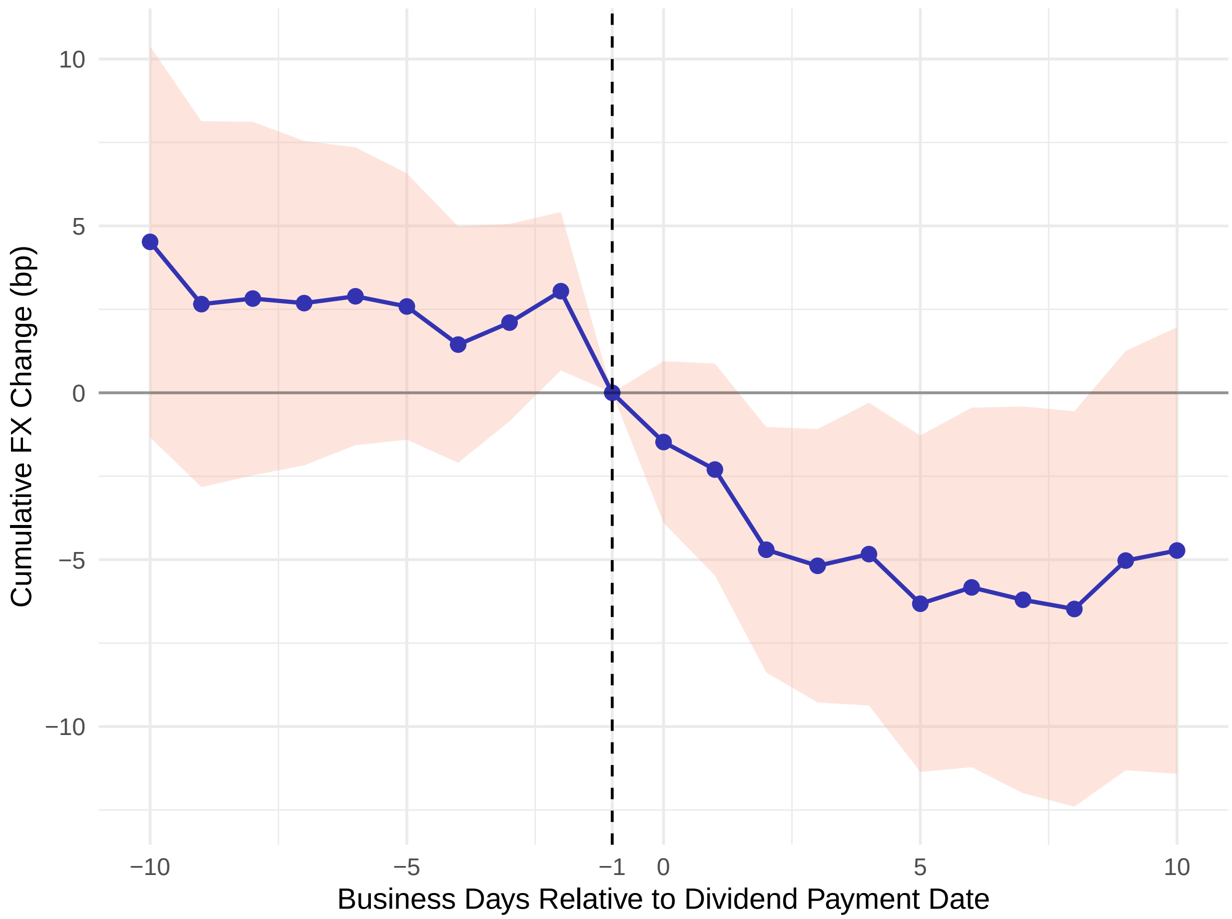


Figure 1.5: Price Impact of Large Dividends on the Foreign Exchange Rate Around Dividend Announcement Dates

This figure shows the coefficients β_h estimated in the baseline regression Eq (1.2), where t is the announcement date instead of the payment date. Dividends are aggregated from the company level to the currency level by the announcement dates. The regression includes controls, the currency fixed effect, and the time fixed effect. The controls include stock market returns and FX implied volatilities. The sample period is from January 2001 to June 2023. The standard errors are two-way clustered at the date level and the currency level

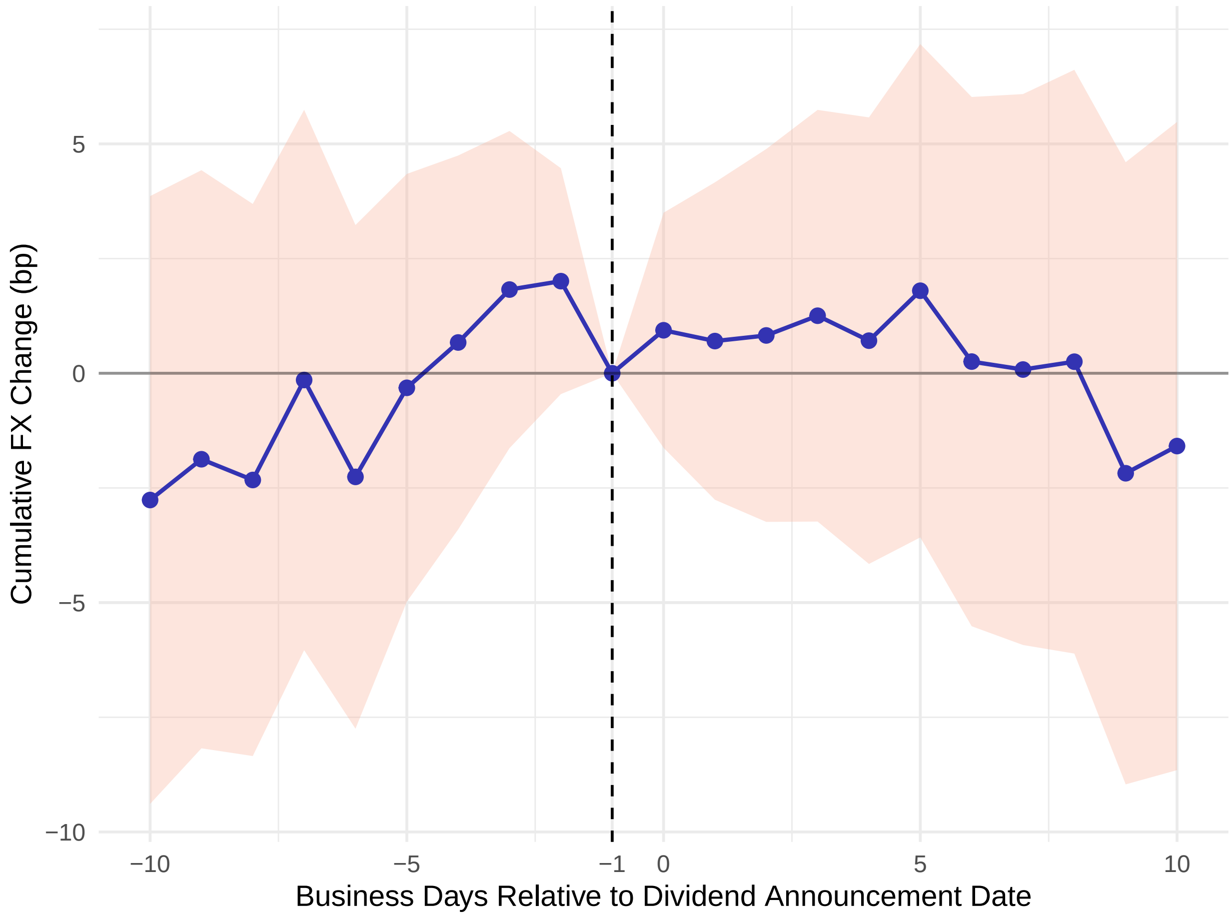


Figure 1.6: Dividend Repatriation Channel - A Case Study

This figure shows the cash position evolution of First Trust Developed Markets ex-US AlphaDEX® Fund (FDT) from November 30, 2022 to December 9, 2022. During this period of time, there are no fund inflows or outflows, no changes in underlying stock positions, and no distributions to the ETF investors. Calculated from FDT's portfolio holdings and the dividend payment information, the fund should receive dividend payments in JPY (orange bar) from its portfolio holdings of Japanese companies from November 30, 2022 (Wednesday) to December 2, 2022 (Friday), with the dividend payment on December 1, 2022 (Thursday) being the largest. In the meantime, dividends received in other currencies are negligible. The JPY dividends appeared on FDT's JPY cash account (red line) on December 5, 2022 (Monday), after which the JPY cash position decreased while the USD cash position (blue line) increased by a similar amount.

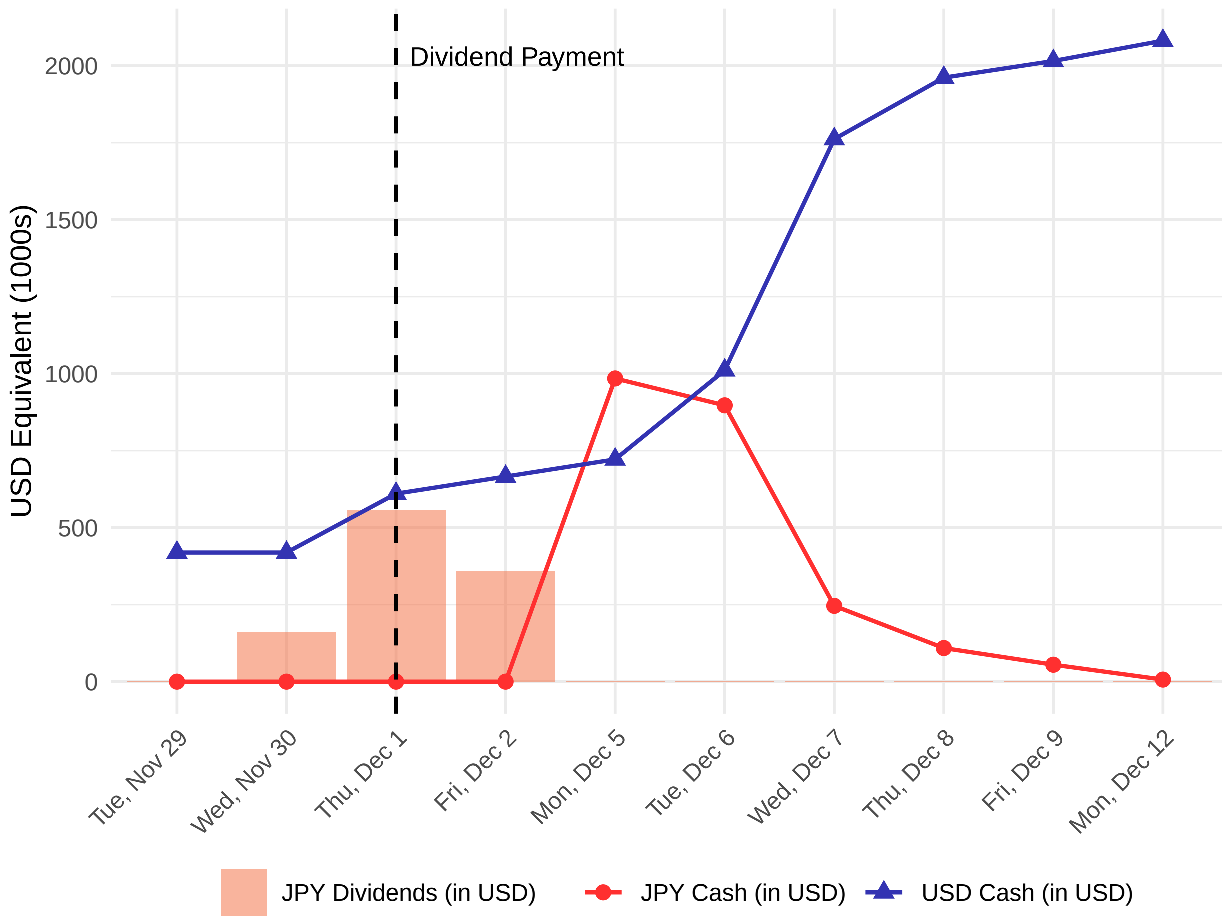


Figure 1.7: Model Timeline and Equilibrium

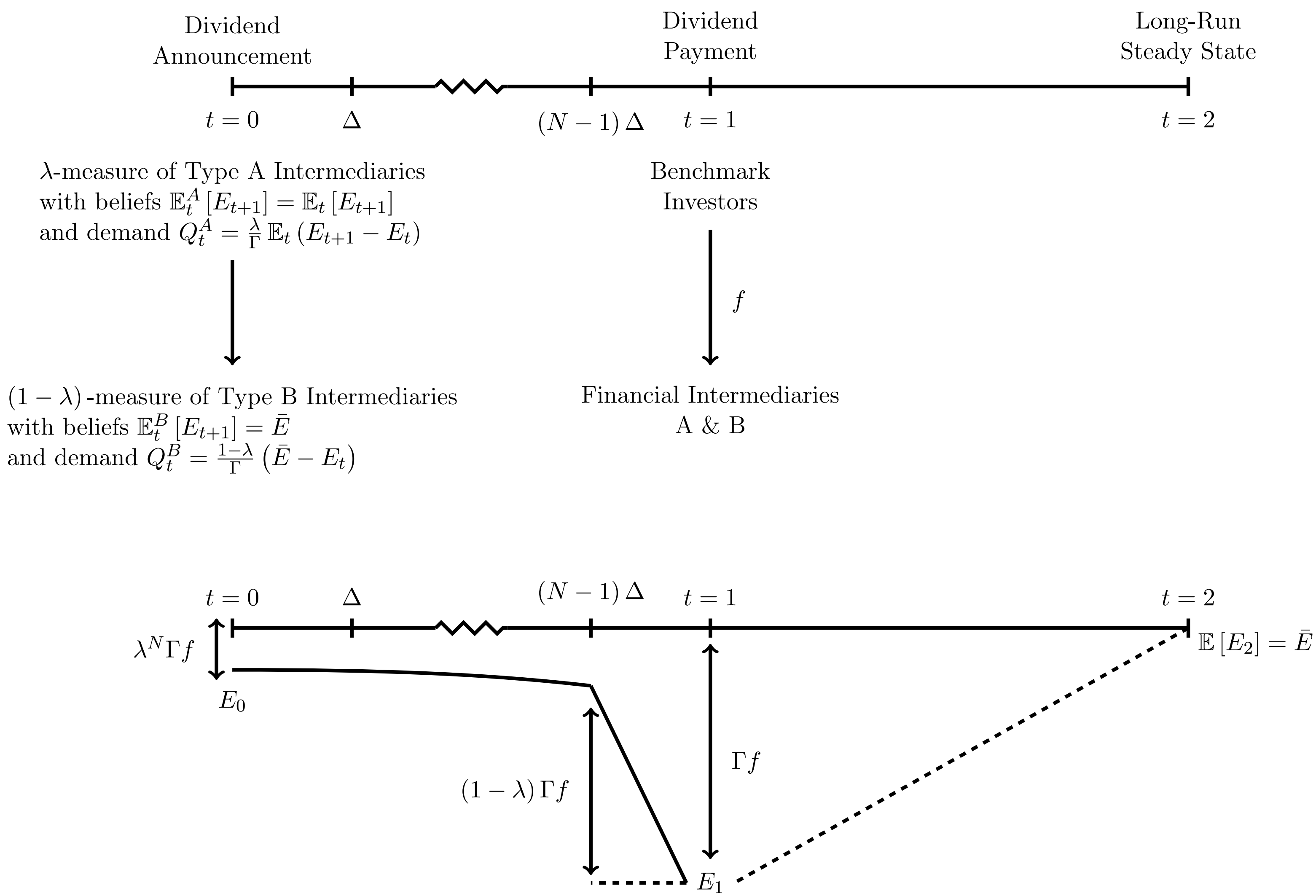


Figure 1.8: Equilibrium FX Dynamics and Positions of Intermediaries

This figure shows the expected value of the exchange rate and positions of intermediaries in equilibrium, according to Eq (1.8)-(1.11). The parameters are $\bar{E} = 1, N = 5, \Delta = 1/5, \lambda = 0.3, \Gamma = 0.01, f = 1$. The calibration of λ is in the main text. Time $t = 0$ is the dividend announcement date and time $t = 1$ is the payment date. A negative change in the exchange rate E means GBP depreciates against USD. Negative Q means a short position in GBP.

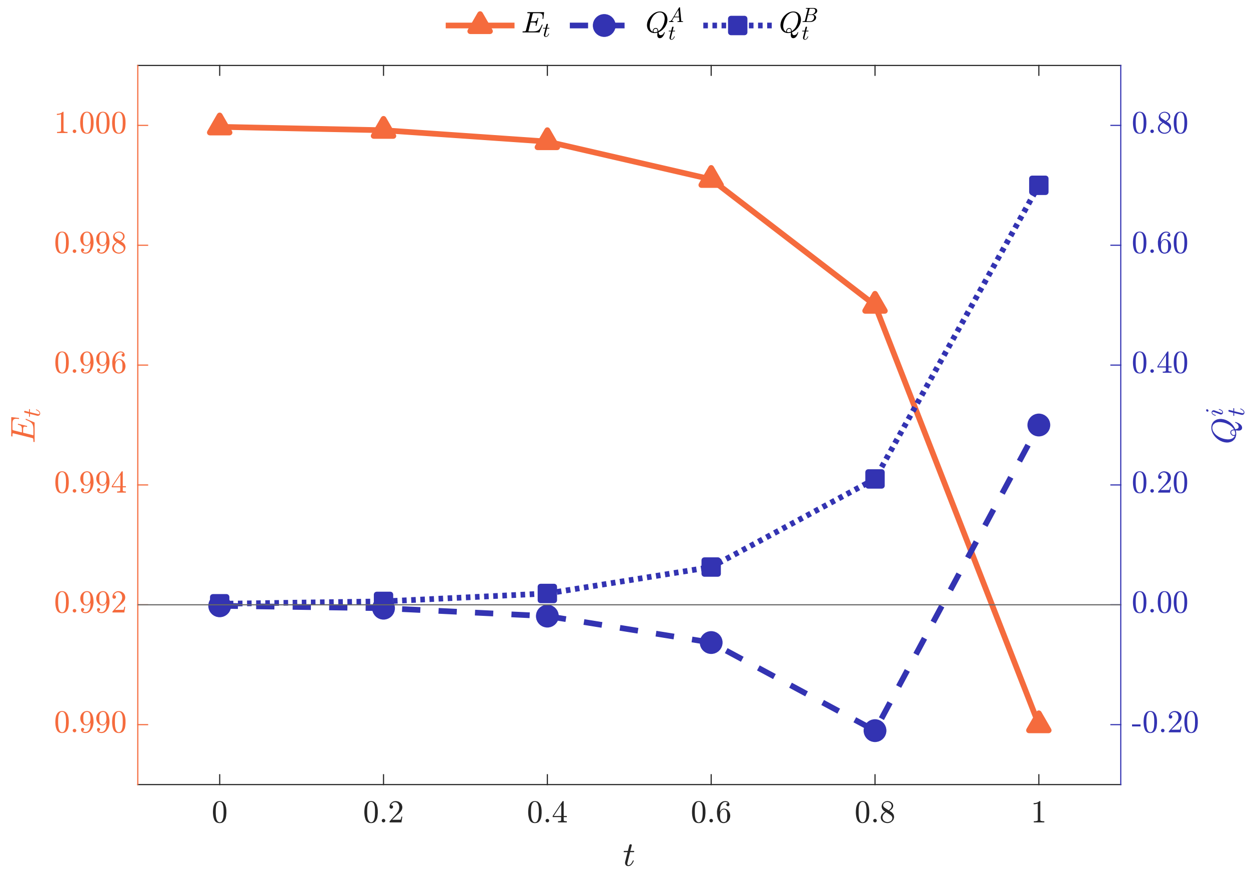


Figure 1.9: Comparison Between Estimates of Dividends to Foreign Investors: Compustat Global/CRSP vs. Balance of Payments

This figure compares the dividends paid out to foreign investors, calculated from Compustat Global/CRSP vs imputed from the Balance of Payments, at an annual frequency in billion USD. Each dot in the figure represents currency-year. For the y-axis, dividends paid out to foreign investors calculated from Compustat Global/CRSP, I first aggregate dividend payments by payment date in each currency area, then I multiply by the foreign ownership calculated imputed from the Balance of Payments. For the x-axis, dividends paid out to foreign investors imputed from the Balance of Payments, I use Dividends on Equity Excluding Investment Fund Shares (BMIPIPED) if the country reports the data item. Otherwise, I use Investment Income on Equity and Investment Fund Shares (BMIPIPE), scaled by the ratio of ILPEEO/ILPE, where ILPEEO represents Equity Other Than Investment Fund Shares, and ILPE represents Equity and Investment Fund Shares, both under Liabilities of Portfolio Investment. See Appendix 1.10.1 for details on the indicators in the Balance of Payments. See Appendix 1.10.2 for details on the calculation of foreign ownership. The sample period is from 2001 to 2022.

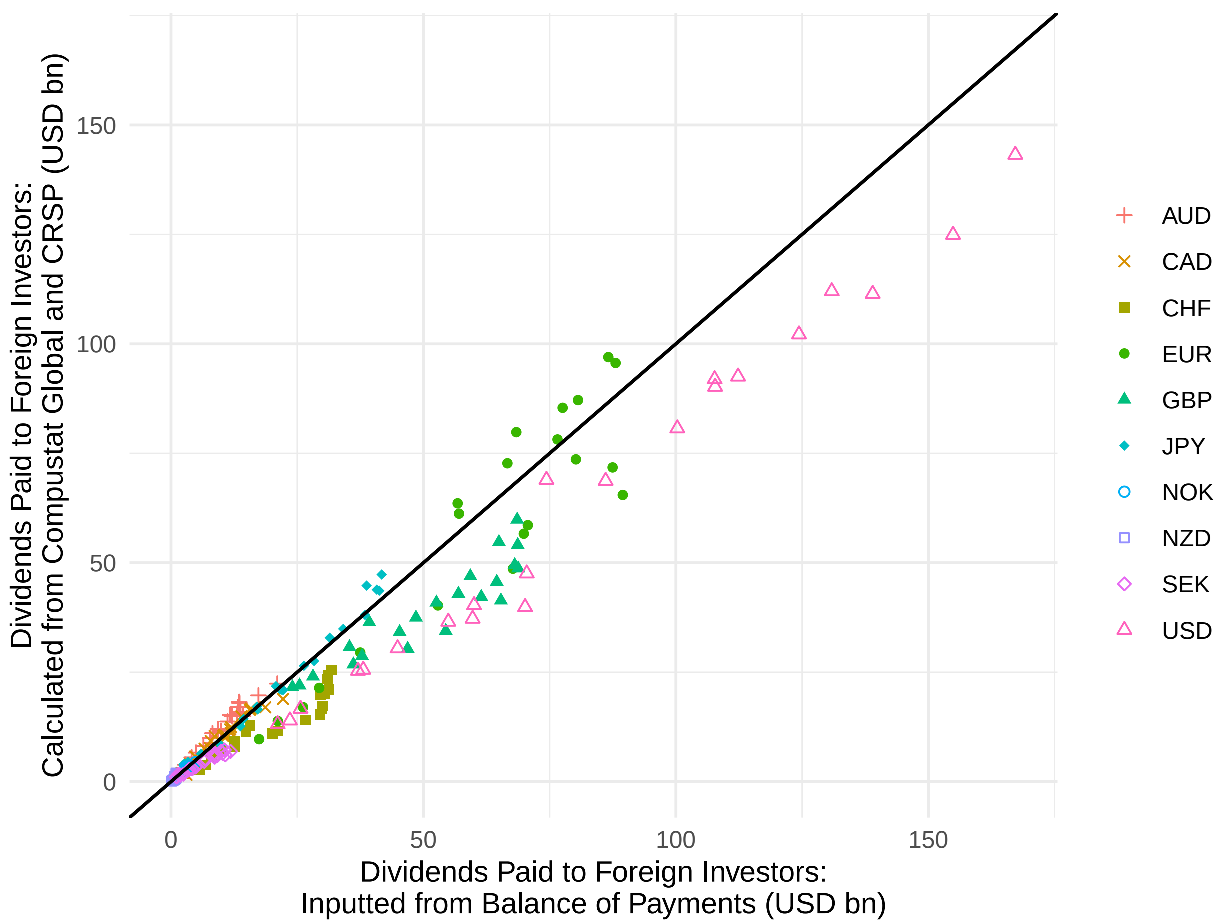


Figure 1.10: Time-Series of FX Regimes

This figure reports the fine classification of FX regimes by Ilzetzki et al. (2019). AUD, EUR, JPY, USD have always been in the freely floating regime, NOK has always been managed floating (anchoring to AUD), and NOK has always been de facto moving band ±2% against Euro. CAD switched from de facto moving band (±2% band against US dollar) to freely floating in June 2002. GBP switched from de facto moving band (±2% band against Euro) to freely floating in January 2009. SEK switched from de facto horizontal band (±2% band against Euro) to de facto moving band (±2% band against Euro) in September 2008. CHF switched to pegging to Euro during September 2011 to January 2015, while in other times, de facto moving band (±2% band against Euro).

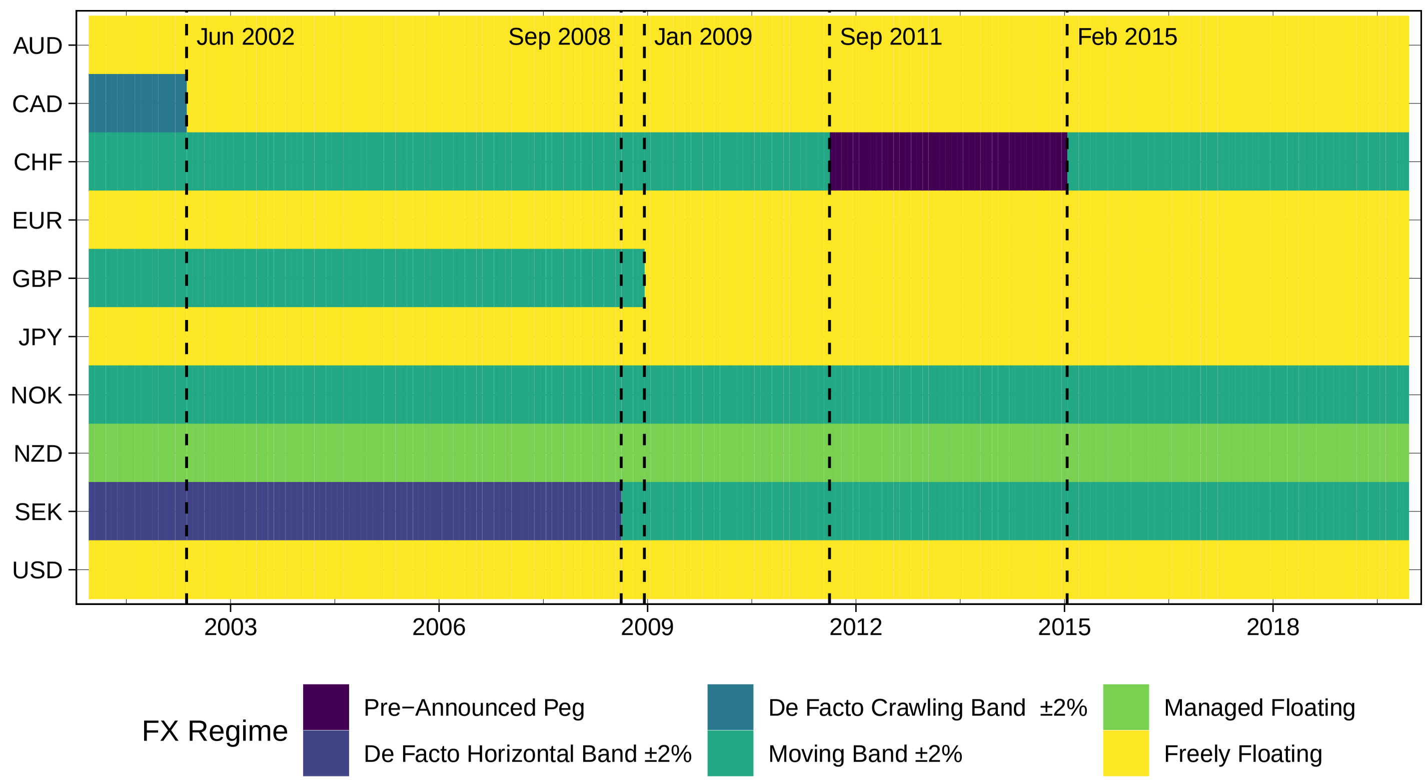


Table 1.1: Stock Market Size and Foreign Ownership

This table provides summary statistics about the size and foreign ownership of stock markets in G10 countries/currency areas, including Australia (AUS), Canada (CAN), Switzerland (CHE), Euro area (EUR), United Kingdom (GBR), Japan (JPN), Norway (NOR), New Zealand (NZL), Sweden (SWE), and the United States (USA). All numbers are the average of annual data from 2001 to 2022. *Stock Market to GDP* is the year-end stock market capitalization divided by nominal GDP, where the market capitalization data is from Bloomberg (after 2003) and the World Bank (before 2023). The nominal GDP is from the World Bank. *Foreign Ownership of Domestic Stock Market* is calculated from the Balance of Payments. Columns *by G10* and *by USA* under *Out of Foreign Ownership* are calculated from the Coordinated Portfolio Investment Survey (CPIS). See Appendix 1.10.2 for more details.

	Stock Market to GDP	Dividends to GDP	Foreign Ownership of Domestic Stock Market	Out of Foreign Ownership	
				by G10	by USA
AUS	1.03	3.7%	28.8%	93.3%	46.5%
CAN	1.16	2.7%	22.4%	96.8%	77.3%
CHE	2.14	3.3%	60.0%	96.9%	48.7%
EUR	0.55	1.5%	32.1%	90.0%	46.0%
GBR	1.19	2.8%	51.7%	89.2%	46.7%
JPN	0.92	1.7%	26.0%	93.5%	52.4%
NOR	0.64	1.8%	26.3%	95.5%	38.3%
NZL	0.36	1.5%	31.9%	96.2%	39.8%
SWE	1.23	2.6%	33.8%	93.7%	35.4%
USA	1.31	2.0%	17.6%	85.1%	-

Table 1.2: Calendar Days Between Dividend Announcement and Payment

This table shows the number of calendar days between the dividend announcement date and the dividend payment date at the firm level across G10 countries/currency areas. Dividend information is released on the dividend announcement date, including dividend size and other dividend-related dates except for Japan. For Japan, I calculate the number of calendar days between the ex-date and the payment date, as companies in Japan typically do not confirm the dividend amount before the ex-date, though the guidance of dividends is usually available almost one year in advance. The sample period is from January 2001 to June 2023.

		Calendar Days			
	Observations	Mean	p25	p50	p75
AUS	17,991	48.3	32	43	58
CAN	55,640	39.4	27	31	44
CHE	2,703	52.1	35	48	63
EUR	35,598	62.9	41	58	83
GBR	32,993	68.0	43	63	84
JPN	106,307	82.6	72	87	93
NOR	2,142	70.3	37	72	97
NZL	2,789	36.4	24	32	44
SWE	4,406	92.7	68	84	97
USA	133,672	43.1	28	37	52
All	394,241	58.0	32	55	80

Table 1.3: Comparison Between Dividend and Portfolio Flows

This table compares average dividend flows with other financial flows and trade flows from the Balance of Payments (BOP) between 2001 and 2022. All numbers are in billions of USD. *Dividends on Equity To Foreign Investors* is investment income on equity and investment fund shares on the debit side (BMIPIPE), while *Dividends on Equity From Foreign Investors* is on the credit side (BXIPIPE). Under *Portfolio Investment of Equity*, *Net Acquisition of Assets* is a country's purchase of foreign countries' equity and investment fund shares (BFPAE), while *Net Incurrence of Liabilities* is foreign countries' purchase of a country's equity and investment fund shares (BFPLE). Similarly, under *Portfolio Investment of Debt*, *Net Acquisition of Assets* is a country's purchase of foreign countries' debt securities (BFPAD), while *Net Incurrence of Liabilities* is foreign countries' purchase of a country's debt securities (BFPLD). *Net Exports* is exports minus imports of goods and services (BGS). See Appendix 1.10.1 for more details on the indicators.

	Dividends on Equity		Portfolio Investment Equity		Portfolio Investment Debt		
	To Foreign Investors	From Foreign Investment	Net Acquisition of Assets	Net Incurrence of Liabilities	Net Acquisition of Assets	Net Incurrence of Liabilities	Net Exports
AUS	12.0	10.6	26.9	13.8	17.9	49.9	6.2
CAN	10.1	13.5	18.9	11.2	14.3	58.4	-2.9
CHE	22.0	12.7	7.9	-2.8	15.4	2.2	50.8
EUR	139.8	62.1	118.8	238.4	289.4	173.2	252.6
GBR	51.0	35.0	-6.8	8.7	37.9	142.3	-45.9
JPN	22.9	36.5	31.2	26.6	98.6	82.4	2.0
NOR	3.8	18.0	26.4	3.1	26.2	14.4	44.9
NZL	0.8	0.9	1.5	0.9	1.6	4.4	0.2
SWE	6.9	9.4	9.6	1.5	8.7	17.7	23.3
USA	98.3	160.7	137.2	91.5	130.5	525.8	-580.5

Table 1.4: FX Cut-off Time and Stock Market Closing Time

This table shows the primary stock market closing time of the regular trading hour in different countries/currency areas. The data is sourced from Bloomberg. The FX cut-off time is the closest hour equal to or immediately after the stock market closing time.

	Time Zone	Stock Market Close	FX Cut-Off Time
AUD	Australia/Sydney	16:00	16:00
CAD	U.S./Eastern	16:00	16:00
CHF	Europe/Zurich	17:20	18:00
EUR	Europe/Paris	17:30	18:00
GBP	Europe/London	16:30	17:00
JPY	Asia/Tokyo	15:00	15:00
NOK	Europe/Oslo	16:20	17:00
NZD	Pacific/Auckland	16:45	17:00
SEK	Europe/Stockholm	17:25	18:00
USD	U.S./Eastern	16:00	16:00

Table 1.5: Performance of Dividend-Based Currency Strategy

This table presents the performance profiles for the dividend-based currency strategy under different parameters, before and after the transaction costs. The transaction cost, i.e., bid-ask spread, is assumed to be 1 basis point for all currencies at all times. The dividend-based currency strategy takes the following form: for each country/currency area k and date t, if in the previous l days, the combined dividend payments in the country k rank in its top p-percentile in the rolling 1-year window, then we sell currency k against USD, and hold the position for one day. If there are several currencies that satisfy this criterion, then the strategy puts $1 on each position. The numbers in the brackets are t-statistics. Alpha, DOL, CAR, MOM, VAL are the coefficients from factor-spanning regression Eq (1.1) at the monthly frequency. The standard errors of the Sharpe ratio are calculated using Lo (2002).

Top $p = 5\%$	Before Transaction Costs			After Transaction Costs		
Lookback Period l	1 Day	2 Days	1 Week	1 Day	2 Days	1 Week
Panel A. Summary Statistics						
Mean	3.0	4.4	3.3	2.3	3.6	2.3
	[2.65]	[3.25]	[2.16]	[2.08]	[2.66]	[1.51]
Sharpe Ratio	0.56	0.68	0.45	0.44	0.56	0.32
	[2.65]	[3.25]	[2.16]	[2.08]	[2.66]	[1.51]
Zero Position Days	0.79	0.75	0.71	0.79	0.75	0.71
Panel B. Regressions $rx = a + b_1 DOL + b_2 CAR + b_3 MOM + b_4 VAL + e_t$						
Alpha	0.22	0.36	0.30	0.17	0.30	0.22
	[2.32]	[3.22]	[2.49]	[1.78]	[2.65]	[1.85]
DOL	0.28	0.41	0.59	0.28	0.41	0.59
	[5.42]	[6.91]	[9.19]	[5.41]	[6.94]	[9.24]
CAR	0.09	0.07	0.05	0.09	0.07	0.04
	[1.64]	[1.11]	[0.68]	[1.59]	[1.03]	[0.58]
MOM	0.01	0.03	0.05	0.01	0.03	0.04
	[0.22]	[0.58]	[0.86]	[0.23]	[0.60]	[0.87]
VAL	-0.07	-0.08	0.10	-0.07	-0.08	0.10
	[-1.57]	[-1.42]	[1.74]	[-1.56]	[-1.40]	[1.79]

Table 1.6: Price Impact of Large Dividends on the Foreign Exchange Rate

This table compares estimates of the price impact of dividends on the foreign exchange rate using different identification strategies. Panel *OLS* reports β_h estimated from Eq (1.2), without controls and fixed effects. Panel *OLS with Controls* controls for stock market returns an FX implied volatilities. Panel *OLS with Controls and Time Fixed Effects* further adds time fixed effects. Panel *Two-Way Fixed Effects with Controls* is the baseline regression Eq (1.2), with controls, time fixed effects, and currency fixed effects. The standard errors are two-way clustered at the date level and the currency level. Panel *Difference-in-Difference* is estimated from the alternative identification strategy in Section 1.10.4, where control group units are equally weighted. Panel *Synthetic Control* is estimated from the alternative identification strategy in Section 1.10.4, where weights on the qualified controls are optimized from Eq (1.23). The sample period is from January 2001 to June 2023.

	Business Days Relative to Dividend Payment Date										
	0	1	2	3	4	5	6	7	8	9	10
OLS											
Coefficients	0.04	-1.77	-5.73**	-6.11**	-5.54*	-7.76**	-7.99**	-9.30**	-9.26**	-7.36	-5.89
Standard Errors	(1.46)	(2.14)	(2.58)	(3.01)	(3.27)	(3.61)	(3.92)	(4.25)	(4.44)	(4.72)	(5.07)
OLS with Controls											
Coefficients	0.13	-1.65	-5.60**	-5.97**	-5.38*	-7.58**	-7.81**	-9.12**	-9.05**	-7.15	-5.67
Standard Errors	(1.45)	(2.12)	(2.56)	(3.00)	(3.26)	(3.60)	(3.91)	(4.24)	(4.42)	(4.70)	(5.05)
OLS with Controls and Time Fixed Effects											
Coefficients	-1.47	-2.29	-4.70**	-5.18**	-4.83**	-6.32**	-5.83**	-6.20**	-6.48**	-5.03	-4.73
Standard Errors	(1.24)	(1.62)	(1.88)	(2.09)	(2.32)	(2.57)	(2.75)	(2.96)	(3.02)	(3.21)	(3.42)
Two-Way Fixed Effects with Controls											
Coefficients	-1.48	-2.30	-4.70**	-5.18**	-4.83**	-6.32**	-5.83**	-6.20**	-6.48**	-5.03	-4.73
Standard Errors	(1.24)	(1.62)	(1.88)	(2.09)	(2.31)	(2.57)	(2.75)	(2.96)	(3.02)	(3.21)	(3.41)
Difference-in-Difference											
Coefficients	-1.81	-4.84**	-6.54***	-7.15**	-8.21***	-9.16**	-10.02**	-10.76**	-12.55***	-11.91***	-11.57**
Standard Errors	(1.47)	(2.14)	(2.48)	(2.82)	(3.16)	(3.56)	(3.91)	(4.26)	(4.34)	(4.50)	(4.65)
Synthetic Controls											
Coefficients	-2.07	-4.57**	-5.76**	-6.22**	-9.10***	-9.72***	-10.52***	-11.83***	-12.51***	-12.45***	-11.99***
Standard Errors	(1.61)	(2.17)	(2.50)	(2.84)	(3.14)	(3.40)	(3.68)	(3.94)	(4.10)	(4.33)	(4.55)

(Continued on the next page)

Table 1.6 (Continued): Price Impact of Large Dividends on the Foreign Exchange Rate

	Business Days Relative to Dividend Payment Date										
	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
OLS											
Coefficients	-2.63	-2.77	-2.64	-2.63	-1.80	-1.50	-1.81	-1.34	0.98	-	0.04
Standard Errors	(4.58)	(4.26)	(4.02)	(3.70)	(3.35)	(2.92)	(2.57)	(2.12)	(1.53)	-	(1.46)
OLS with Controls											
Coefficients	-2.37	-2.54	-2.42	-2.42	-1.62	-1.37	-1.69	-1.26	1.03	-	0.13
Standard Errors	(4.59)	(4.26)	(4.03)	(3.71)	(3.34)	(2.92)	(2.57)	(2.11)	(1.53)	-	(1.45)
OLS with Controls and Time Fixed Effects											
Coefficients	4.48	2.62	2.79	2.66	2.87	2.57	1.43	2.09	3.04**	-	-1.47
Standard Errors	(2.99)	(2.80)	(2.71)	(2.49)	(2.28)	(2.04)	(1.81)	(1.51)	(1.21)	-	(1.24)
Two-Way Fixed Effects with Controls											
Coefficients	4.52	2.66	2.82	2.69	2.89	2.59	1.44	2.10	3.04**	-	-1.48
Standard Errors	(2.99)	(2.80)	(2.70)	(2.48)	(2.28)	(2.04)	(1.81)	(1.51)	(1.21)	-	(1.24)
Difference-in-Difference											
Coefficients	0.58	1.87	3.00	2.39	4.99	4.52	3.67	4.03*	2.05	-	-1.81
Standard Errors	(4.32)	(4.08)	(3.85)	(3.51)	(3.19)	(2.85)	(2.50)	(2.08)	(1.54)	-	(1.47)
Synthetic Controls											
Coefficients	4.35	4.86	5.06	2.62	4.32	4.20	3.31	3.29*	1.18	-	-2.07
Standard Errors	(3.85)	(3.61)	(3.43)	(3.21)	(2.96)	(2.66)	(2.29)	(1.84)	(1.36)	-	(1.61)

(Continued from the previous page)

Table 1.7: Time Variation of Price Impact of Dividend Flows on FX Rates

This table reports the price impact coefficients of dividends paid out to foreign investors on the foreign exchange rate in Eq(1.15). The variable $DivOut_{k,t}$ is country k's (normalized) dividends paid out to foreign investors on date t, calculated using total dividend payments from Compustat Global/CRSP multiplied by the foreign ownership, then normalized by the previous year-end its stock market capitalization, both in local currency. The controls include stock market returns and FX implied volatilities. Columns 1-4 are on different subsamples. For regressions with interaction terms with subsample indicators, the fixed effects and controls are fully saturated. Column 5 reports results in Eq (1.16). The standard errors are two-way clustered at the date level and the currency level. The sample period is from January 2001 to June 2023.

Panel A. Intermediary Capital Ratio					
	(1)	(2)	(3)	(4)	(5)
$\Delta_2 e^{US/LC}_{t+2}$	All	CR $\geq$ p50	CR $<$ p50	All	All
$DivOut_{i,t}$	-0.806***	-0.192	-1.209***	-0.192	-2.123**
	(0.259)	(0.348)	(0.363)	(0.348)	(0.845)
$\mathbb{1}\{\text{CR} < \text{p50}\} \times DivOut_{i,t}$				-1.018**	
				(0.503)	
CR $\times$ $DivOut_{i,t}$					20.513*
					(11.643)
Observations	50463	25245	25218	50463	50463
Adjusted R^2	0.518	0.522	0.516	0.518	0.518

Panel B. CIP Deviation				
	(1)	(2)	(3)	(4)
$\Delta_2 e^{US/LC}_{t+2}$	All	$\lvert$CIP$\rvert$ $<$ p50	$\lvert$CIP$\rvert$ $\geq$ p50	All
$DivOut_{i,t}$	-0.806***	-0.302	-1.259***	-0.302
	(0.259)	(0.360)	(0.417)	(0.360)
$\mathbb{1}\{\lvert\text{CIP}\rvert \geq \text{p50}\} \times DivOut_{i,t}$				-0.957*
				(0.555)
Observations	50463	24290	24749	49039
Adjusted R^2	0.518	0.568	0.552	0.558

Panel C. Currency Implied Volatility				
	(1)	(2)	(3)	(4)
$\Delta_2 e^{US/LC}_{t+2}$	All	IV $<$ p50	IV $\geq$ p50	All
$DivOut_{i,t}$	-0.806***	-0.359	-1.290***	-0.359
	(0.259)	(0.311)	(0.429)	(0.311)
$\mathbb{1}\{\text{IV} \geq \text{p50}\} \times DivOut_{i,t}$				-0.931*
				(0.531)
Observations	50463	24797	24425	49222
Adjusted R^2	0.518	0.538	0.548	0.545

Table 1.8: Comparison Among Estimates of FX Semi-Multipliers

This table compares my estimates using dividend flows with estimates in the existing literature, which are converted into semi-multiplier, i.e., the dollar value of capital flows that can move the exchange rate by 1%. See the footnotes in the main text for details of the conversion.

	Semi-Multipler		Sample	
	Methodology	Est	Currencies	Freq
This paper	Dividend flows	8.1	G10	D
Camanho et al.	GIV on MF rebalancing flows	7.1	USD, EUR, GBP, CAD	Q
Hau et al.	MSCI Global Equity Index redefinition	2.6	33 DM & EM	D
Evans-Lyons	Order flows	1.9	DEM	D
Pandolfi-Williams	Cap 10% in GBI-EM Global Diversified Index induced rebalancing	1.4	16 EM	D
Broner et al.	Addition to WGBI & GBI-EM	5.0	6 EM	D
Aldunate et al.	Chilean FyF induced rebalancing	1.4	CLP	D

Table 1.9: Price Impact by Foreign Exchange Regime

This table reports the price impact coefficients of dividends paid out to foreign investors on the foreign exchange rate in Eq(1.15) under different FX regimes. FX regimes are the fine classifications from Ilzetzki et al. (2019). The variable $DivOut_{k,t}$ is country k's (normalized) dividends paid out to foreign investors on date t, calculated using total dividend payments from Compustat Global/CRSP multiplied by the foreign ownership, then normalized by the previous year-end its stock market capitalization, both in local currency. The controls include stock market returns and FX implied volatilities. The standard errors are two-way clustered at the date level and the currency level. The sample period is from January 2001 to June 2023.

$\Delta_2 e_{t+2}^{US/LC}$	(1) All	(2) Non-Freely Floating	(3) Freely Floating	(4) All
$DivOut_{i,t}$	-0.806***	-0.353	-1.689***	-0.353
	(0.259)	(0.335)	(0.644)	(0.335)
$\mathbb{1}\{\text{FreeFloat}\} \times DivOut_{i,t}$				-1.336*
				(0.721)
Observations	50463	24364	26099	50463
Adjusted R^2	0.518	0.645	0.470	0.567

1.10 Appendix

1.10.1 Data Items From the Balance of Payments

In this paper, I use the following data from the Balance of Payment. Below are the list of indicator codes and indicator names. Note all countries report all data items below. The more detailed the data items, the less likely a country is reporting it.

Dividends on Equity

- Paid to foreign investors
  - BMIPIPE: Current Account, Primary Income, Investment Income, Portfolio Investment, Investment Income on Equity and Investment Fund Shares, Debit
  - BMIPIPED: Current Account, Primary Income, Investment Income, Portfolio Investment, Investment Income on Equity and Investment Fund Shares, Dividends on Equity Excluding Investment Fund Shares, Debit
- Received from foreign investments
  - BXIPIPE: Current Account, Primary Income, Investment Income, Portfolio Investment, Investment Income on Equity and Investment Fund Shares, Credit
  - BXIPIPED: Current Account, Primary Income, Investment Income, Portfolio Investment, Investment Income on Equity and Investment Fund Shares, Dividends on Equity Excluding Investment Fund Shares, Credit

Portfolio Investment

- Asset: investment in foreign countries
  - IAPE: Assets, Portfolio Investment, Equity and Investment Fund Shares
  - IAPEEO: Assets, Portfolio Investment, Equity and Investment Fund Shares, Equity Securities Other Than Investment Fund Shares

    - BFPAE: Financial Account, Portfolio Investment, Net Acquisition of Financial Assets, Equity and Investment Fund Shares
    - BFPAD: Financial Account, Portfolio Investment, Net Acquisition of Financial Assets, Debt Securities

- Liability: investment by foreign countries

    - IAPE: Assets, Portfolio Investment, Equity and Investment Fund Shares
    - IAPEEO: Assets, Portfolio Investment, Equity and Investment Fund Shares, Equity Securities Other Than Investment Fund Shares
    - BFPLE: Financial Account, Portfolio Investment, Net Incurrence of Liabilities, Equity and Investment Fund Shares
    - BFPLD: Financial Account, Portfolio Investment, Net Incurrence of Liabilities, Debt Securities

Others

- BGS: Current Account, Goods and Services, Net

1.10.2 Calculation of Foreign Ownership

In this section, I provide further details on the imputation of foreign ownership underlying Table 1.1 and the construction of $DivOut_{k,t}$ in Eq (1.15).

Foreign ownership is calculated by external liabilities of equity securities other than investment fund shares in portfolio investment (ILPEEO) divided by the stock market capitalization. If the country does not report ILPEEO in the Balance of Payments (BOP), I impute it from external liabilities of equity and investment fund shares in portfolio investment (ILPE) scaled by the backfilled ILPEEO/ILPE ratio. Backfilled ILPEEO/ILPE ratio fills the missing values by the last non-missing values. If ILPEEO is missing throughout the sample, I use ILPE instead. In most countries, ILPEEO/ILPE ratio is high. The major exception is Eurozone, where on average ILPEEO/ILPE ratio is 42%.

The stock market capitalization data is from Bloomberg (after 2003) and the World Bank (before 2023). The Bloomberg market capitalization is calculated from all shares outstanding. It does not include ETFs and ADRs as they do not directly represent companies. Also, it includes only actively traded, primary securities on the country's exchanges to avoid double counting. For years before 2003, I use data from the World Bank.

For the breakdown of foreign ownership into by G10 and by USA, I use data from the Coordinated Portfolio Investment Survey (CPIS). CPIS has bilateral equity holdings data, from which I can calculate a country's external liabilities of equity by other G10 countries and by USA. Note CPIS equity holdings include both equity and investment fund shares, hence it is similar to ILPE in terms of concept. In cases where external equity liabilities aggregated from bilateral equity holdings in CPIS is larger ILPE reported in BOP, I scale down CPIS equity holdings proportionally. The foreign ownership of the stock market by other G10 countries is calculated from foreign ownership calculated in BOP, scaled by the ratio of equity held by other G10 (from CPIS) and ILPE (from BOP). The foreign ownership of the stock market by the USA is calculated similarly.

1.10.3 Proofs

Proof. At the payment date time 1, demand for both intermediaries is

$$Q_1^A = \frac{\lambda}{\Gamma}\mathbb{E}_1[E_2 - E_1] = \frac{\lambda}{\Gamma}(\bar{E} - E_1), \quad Q_1^B = \frac{1-\lambda}{\Gamma}(\bar{E} - E_1) \tag{1.17}$$

The GBP market clearing condition on the payment date $t = 1$ is

$$Q_1^A + Q_1^B - f + \eta_1 = 0 \tag{1.18}$$

where $-f$ is the benchmark investor's selling GBP to repatriate a certain proportion of dividends out of GBP, and η_1 is the noise trader's demand for GBP. Plug in the demand curves for both types of intermediaries, Eq (1.5) and Eq (1.7), we have

$$\lambda\mathbb{E}_1[E_2] + (1-\lambda)\bar{E} - E_1 = \Gamma f + \Gamma(-\eta_1) \tag{1.19}$$

Plug Eq (1.3) into Eq (1.19), we have the exchange rate on the payment date:

$$E_1 = \bar{E} - \Gamma f + \Gamma\eta_1$$

Plug this back into Eq (1.17), we calculate the positions of GBP for both types of intermediaries

$$Q_1^A = \lambda(f - \eta_1), \quad Q_1^B = (1-\lambda)(f - \eta_1)$$

For exchange rates before the payment date, I use backward induction to solve the E_{t_n}, where $t_n = n\Delta, n = 0, ..., N-1$ and $N\Delta = 1$. For simplicity of notation, assume $Var_{t_n}[E_{t_{n+1}}] = \sigma_E^2$. For this to hold, we need the parameter assumption that $\sigma_\eta = 1/(\gamma\sigma_E)$.

The demand for GBP from both type A and type B intermediaries is

$$Q_{t_n}^A = \frac{\lambda}{\Gamma}\mathbb{E}_{t_n}[E_{t_{n+1}} - E_{t_n}], \quad Q_{t_n}^B = \frac{1-\lambda}{\Gamma}(\bar{E} - E_{t_n}) \tag{1.20}$$

Plug into the GBP market clearing at time t_n

$$Q_{t_n}^A + Q_{t_n}^B + \eta_{t_n} = 0 \tag{1.21}$$

we have

$$\lambda\mathbb{E}_{t_n}[E_{t_{n+1}} - E_{t_n}] + (1-\lambda)(\bar{E} - E_{t_n}) + \Gamma\eta_{t_n} = 0$$

which gives

$$E_{t_n} = (1-\lambda)\bar{E} + \lambda \mathbb{E}_{t_n}[E_{t_{n+1}}] + \Gamma \eta_{t_n}$$

Iterated forward, we have

$$\begin{aligned}
E_{t_n} &= (1-\lambda)\bar{E} + \lambda \mathbb{E}_{t_n}[E_{t_{n+1}}] + \Gamma \eta_{t_n} \\
&= (1-\lambda)\bar{E} + \lambda \mathbb{E}_{t_n}[(1-\lambda)\bar{E} + \lambda \mathbb{E}_{t_{n+1}}[E_{t_{n+2}}] + \Gamma \eta_{t_{n+1}}] + \Gamma \eta_{t_n} \\
&= (1-\lambda)\bar{E}(1+\lambda) + \lambda^2 \mathbb{E}_{t_n}[E_{t_{n+2}}] + \Gamma \eta_{t_n} \\
&= (1-\lambda)\bar{E}(1+\lambda+...+\lambda^{k-1}) + \lambda^k \mathbb{E}_{t_n}[E_{t_{n+k}}] + \Gamma \eta_{t_n} \\
&= (1-\lambda)\bar{E}(1+\lambda+...+\lambda^{N-n-1}) + \lambda^{N-n} \mathbb{E}_{t_n}[E_{t_N}] + \Gamma \eta_{t_n} \\
&= (1-\lambda)\bar{E}\frac{1-\lambda^{N-n}}{1-\lambda} + \lambda^{N-n}(\bar{E} - \Gamma f) + \Gamma \eta_{t_n} \\
&= \bar{E} - \lambda^{N-n}\Gamma f + \Gamma \eta_{t_n}
\end{aligned}$$

Plug this exchange rate dynamics back into Eq (1.20), we solve for the demand of both intermediaries at times before the payment date:

$$Q^A_{t_n} = \frac{\lambda}{\Gamma}\mathbb{E}_{t_n}[-\lambda^{N-n-1}\Gamma f + \Gamma \eta_{t_{n+1}} + \lambda^{N-n}\Gamma f - \Gamma \eta_{t_n}] = -\lambda^{N-n}(1-\lambda) f - \lambda \eta_{t_n}$$

$$Q^B_{t_n} = \frac{1-\lambda}{\Gamma}\left(\bar{E} - (\bar{E} - \lambda^{N-n}\Gamma f + \Gamma \eta_{t_n})\right) = (1-\lambda)\lambda^{N-n} f - (1-\lambda)\eta_{t_n}$$

Lastly, for the volatility of the next-period exchange rate to be constant at σ_E, we simply need the parameter assumption that $\sigma_\eta = 1/(\gamma \sigma_E)$, as

$$Var_{t_n}[E_{t_{n+1}}] = \Gamma^2 \sigma^2_\eta = (\gamma \sigma^2_E)^2 \sigma^2_\eta = \sigma^2_E$$

This completes the proof of Proposition 1.

The proof of Proposition 2 is straight-forward.

By the definition of the payment date effect,

$$\mathbb{E}[E_1 - E_{t_{N-1}}] = (\bar{E} - \Gamma f) - (\bar{E} - \lambda \Gamma f) = -(1-\lambda)\Gamma f$$

By the definition of the anticipation effect,

$$\mathbb{E}[E_{t_{N-1}} - \bar{E}] = \mathbb{E}[(\bar{E} - \lambda \Gamma f + \Gamma \eta_{t_{N-1}}) - \bar{E}] = -\lambda \Gamma f \tag{1.22}$$

By the definition of the announcement date effect,

$$\mathbb{E}[E_0 - \bar{E}] = (\bar{E} - \lambda^N \Gamma f) - \bar{E} = -\lambda^N \Gamma f$$

This completes the proof of Proposition 2.

$\square$

1.10.4 Additional Identification Strategies

The baseline identification strategy in Section 1.5 assumes that unspecified time-varying confounding has the same effect on all currencies and hence can be absorbed by the time effect. However, different currencies may have heterogeneous loadings on the underlying factor. For example, the commodity price increase may benefit commodity-exporting countries' terms of trade and currencies. In addition, instead of being constant, the effect of the underlying confounding factors may be time-varying. Below, I develop alternative strategies to confirm that the baseline results are robust under various identification strategies, i.e., the foreign exchange rate depreciates shortly after the dividend payment dates, while the anticipation effect before the payment date is limited.

Difference-in-Difference

In this section, I present the results estimated by difference-in-difference (DiD). This is a special case of the synthetic controls in Section 1.10.4 in the sense that DiD puts equal weights on the control group currencies. See Section 1.10.4 for the definitions of the treated currency and the control group currencies. The standard errors are two-way clustered at currency level and date level. Figure 1.13 shows the results, which confirm the same pattern as in Section 1.5.1.

I also apply the method to each currency individually. Figure 1.15 shows the price impact estimates for each G10 currency against USD. When estimated individually, for many currencies we do not have enough power. Nevertheless, the point estimates suggest that the patterns of depreciation pressure after the dividend payment date are present for most currencies.

Synthetic Controls

In this section, I develop an alternative identification strategy using the idea of synthetic control (e.g., Abadie (2021)), which carefully chooses a linear combination of control group currencies that best replicates the movements of the treated currency. By taking the difference between the treated currency and this linear combination, one can take out the unspecified confounding variables in a flexible way. In addition, as taking the difference absorbs the noisy variation in the estimation, this method results in a more precise estimate,

Specifically, I define a dividend event as a currency-day pair (k_0, t_0) where the country k_0 has a top 5% largest dividend within the currency-year on the payment date t_0. Denote the event date by $t = 0$ and all days relative to it are in trading days. One concern of the discretization of dividend indicator $\mathbb{D}_{k,t}$ is that dividend payments immediately below the size threshold are classified as nonevents, which may pollute the comparison of the treated and the controls. To address this concern, I incorporate a buffer in defining the control group units, i.e., instead of the top 5% size threshold when defining treated currencies, the controls are currencies that do not have the top 10% largest dividend payments within currency-year over the event window, from -10 days to +10 days. The results are robust to both choices of size threshold and buffers.

Among the control group currencies $\mathcal{C}$, I randomly select one p_0 as the placebo. Denote the remaining control group currencies as $\mathcal{C}'$. I find non-negative weights $\{w_k\}_{k\in\mathcal{C}'}$ that sum up to 1, and the linear combination of currencies best tracks the movement of treated currency k_0 over the estimation window [-70,-11]. In other words, the synthetic control weights are calculated from the following optimization problem:

$$\min_{\{w_k\}_{k\in\mathcal{C}'}} \sum_{t=-70}^{-11} |\Delta \ln E_{k_0,t}^{US/LC} - \sum_{k\in\mathcal{C}'} w_k \Delta \ln E_{k,t}^{US/LC}|^2 \tag{1.23}$$

s.t.

$$\sum_k w_k = 1, w_k \geq 0, \forall k \in \mathcal{C}'$$

where the foreign exchange rates are snapshots at the local stock market closing time of the treated currency k_0. With estimated weights, I compare the cumulative FX movement of the treated currency with the synthetic control over the event window [-10,10] for the dividend event (k_0, t_0), where I normalize the pre-event t=-1 to be 0. The treatment effect is as follows:

$$\Delta_h e_{k_0,t} - \sum_{k\in\mathcal{C}'} w_k \Delta_h e_{k,t}, \quad h = -10, ..., 0, ..., 10 \tag{1.24}$$

where $\Delta_h e_{k,t} = \ln E_{k,t+h}^{US/LC} - \ln E_{k,t-1}^{US/LC}$ is the h-day cumulative log change of the foreign exchange rate. The placebo effect for this event is calculated similarly, with the synthetic control weights optimized for the placebo itself using the same procedure as in Eq (1.23). The foreign exchange rates involved are cut at the local stock market closing time of the placebo currency p_0.

$$\Delta_h e_{p_0,t} - \sum_{k\in\mathcal{C}'} w_k^{(p_0)} \Delta_h e_{k,t}, \quad h = -10, ..., 0, ..., 10 \tag{1.25}$$

The average treatment effect (ATT) is the average of Eq (1.24) across all events. The standard errors are calculated from the placebo effect in Eq (1.25) across all events.

Figure 1.12 illustrates how this method works. August 5, 2022 is a dividend event date ($t = 0$) for the UK, as it has a large dividend payment of ≈1.9 billion GBP on this date, among which 1.1 billion is Vodafone's dividend.[49] Over the event window [-10,10] trading days, the qualified controls include AUD, CHF, EUR, JPY, NOK, NZD, SEK, as their countries do not have top 10% dividend payments over the trading days $t = -10$ to $t = +10$. As SEK is selected as the placebo randomly, the remaining control group $\mathcal{C}'$ includes AUD, CHF, EUR, JPY, NOK, NZD. Solving the optimization problem (1.23) gives the following best mimicking linear combination over the estimation window from $t = -70$ to $t = -11$: 15.7% AUD + 15.0% CHF + 30.6% EUR + 14.3% JPY + 9.0% NOK + 15.4% NZD. As Figure 1.12 shows, the synthetic control tracks the day-to-day movement of the treated currency well during the estimation window. The underlying identification assumption is that going forward into the event window, the synthetic control captures the unspecified confounding factors in a flexible way.

Figure 1.14 Panel A shows the average treatment effect. It confirms the pattern in Section 1.5.1. Upon and after the country's large dividend payment dates, the local currency starts to depreciate against USD. The price effect of exchange rates before the dividend payment, i.e., the anticipation effect, is limited and statistically insignificant. In contrast, a placebo currency does not have large dividend payments during the event window. Therefore,

49. For the financial year ending 31 March 2017 and beyond, Vodafone's dividends have been declared in EUR and paid in Euro, GBP and USD. See `https://investors.vodafone.com/individual-shareholders/dividends`

there should be no depreciation pressure on its exchange rate. Figure 1.14 Panel B confirms this is indeed the case.

Figure 1.16 shows the estimates applies the difference-in-difference methodology in Section 1.10.4 to each currency pair.

One concern of using the synthetic control or DiD is the violation of Stable Unit Treatment Values Assumption (SUTVA) assumption. Repatriation of the treated currencies to the control group currencies may cause control group currencies to appreciate against US dollars. Moreover, different foreign exchange rates influence each other through general equilibrium forces. To address the spillover concern, I conduct regression analysis to ensure the spillover effect is small. Specifically, I run the following regression:

$$\ln E_{k,t+h}^{US/LC} - \ln E_{k,t-1}^{US/LC} = \alpha_h + \beta_h \mathbb{D}_{k,t} + \gamma_h \mathbb{D}_{-k,t} + Controls + \gamma_k^{(h)} + \epsilon_{k,t+h} \qquad (1.26)$$

where the indicator $\mathbb{D}_{k,t} = 1$ if country k has a large dividend payment on date t, while indicator $\mathbb{D}_{-k,t} = 1$ if any other country has a large dividend payment. As the time fixed effect will absorb $\mathbb{D}_{-k,t}$, I only include the currency fixed effect in Eq (1.26). As before, the controls include stock market returns and FX implied volatilities. Table 1.11 reports own-effect β_h and cross-effect γ_h. As we can see, β_h estimated is similar to Table 1.7. In the meantime, the cross-effect γ_h, i.e., other countries' dividend payment on country k's exchange rate against USD is insignificant.

1.10.5 Additional Results

Price Impact of Dividends on FX: Pre-GFC vs. Post-GFC

Figure 1.17 compares before and after the GFC. This figure compares the coefficients β_h estimated by Eq (1.2) in the subsample before and after the 2007–2008 Global Financial Crisis (GFC). I define the pre-GFC subsample as before December 2007, and the post-GFC subsample as after June 2009, inclusive.[50] As the point estimates indicate, the local currency depreciates more against USD after the country's large dividend payments in the post-GFC subsample.[51] For example, two days after a country's large dividend payment, its currency depreciates 7.4 basis points vis-à-vis USD in the post-GFC period on average, while before the financial crisis, it only depreciates 1.5 basis points.

From the lens of the model, there are two reasons for the increase in the price impact of dividend payments on the foreign exchange rate. On the one side, with the development of financial integration and passive investing, there is a substantial increase of foreign ownership by benchmark investors like ETFs and mutual funds, which makes the dividend repatriation channel stronger. That is to say, for the same amount of dividend payments in local currency, the dividend repatriation flows f out of this currency is larger. In fact, as Figure 1.1 shows, average across the other G10 countries, the market value of US-domiciled ETFs' holdings as a percentage of the local stock market capitalization grows from 0.7% in 2011 to 3.2% in 2020, more than quadruple in 9 years. Meanwhile, US-domiciled mutual funds grow from 1.93% in 2002 to 4.6% in 2011 to 6.6% in 2020. On the other side, after the 2007-2008 financial crisis, more stringent regulations on financial intermediaries have made their balance sheet constraints tighter. Therefore, financial intermediaries need more compensation to bear the same amount of risk, i.e., Γ increase.

50. Per NBER business cycle dating, the peak of the financial crisis is December 2007, and the trough month is June 2009.

51. The standard errors in the pre-GFC subperiod are too large to conclude the differences are statistically significant.

Table 1.10: G10 Currencies and Stock Markets

This table lists the abbreviations for G10 currencies and their countries/currency areas used in the paper. It also lists their primary stock market index from which I calculate the local stock market returns.

Country/Currency Area		Currency		Stock Market Index	
Abbrv	Name	Abbrv	Name	Abbrv	Name
AUS	Australia	AUD	Australian dollar	ASX200	S&P/ASX 200 Index
CAN	Canada	CAD	Canadian dollar	TSX	S&P/TSX Composite Index
CHE	Switzerland	CHF	Swiss franc	SMI	Swiss Market Index
EUR	European Union	EUR	Euro	SX5E	Euro Stoxx 50
GBR	United Kingdom	GBP	Pound sterling	FTSE100	FTSE 100 Index
JPN	Japan	JPY	Japanese yen	NI225	Nikkei 225 Index
NOR	Norway	NOK	Norwegian krone	OBX25	OBX Stock Index
NZL	New Zealand	NZD	New Zealand dollar	NZX50	S&P/NZX 50 Index
SWE	Sweden	SEK	Swedish krona	OMXS30	OMX Stockholm 30 Index
USA	United States	USD	U.S. dollar	SP500	S&P 500 Index

Table 1.11: Own-Effect and Cross-Effect of Dividends on Exchange Rates

This table reports the coefficients β_h and γ_h in Eq (1.26).

	Days Relative to Dividend Payment Date										
	0	1	2	3	4	5	6	7	8	9	10
No FE											
Coefficients β	-0.057	-1.470	-5.046**	-5.410**	-4.958*	-7.263**	-7.490**	-8.666**	-8.794**	-7.060*	-5.781
Standard Errors	(1.352)	(1.941)	(2.323)	(2.705)	(2.933)	(3.237)	(3.506)	(3.802)	(3.944)	(4.187)	(4.505)
Coefficients γ	0.848	-0.817	-2.510	-2.547	-1.925	-1.457	-1.471	-2.061	-1.166	-0.421	0.504
Standard Errors	(1.153)	(1.751)	(2.189)	(2.554)	(2.871)	(3.138)	(3.383)	(3.637)	(3.895)	(4.123)	(4.338)
Currency FE											
Coefficients β	-0.059	-1.471	-5.047**	-5.409**	-4.956*	-7.260**	-7.487**	-8.665**	-8.792**	-7.056*	-5.777
Standard Errors	(1.352)	(1.941)	(2.324)	(2.705)	(2.932)	(3.236)	(3.504)	(3.800)	(3.943)	(4.186)	(4.503)
Coefficients γ	0.839	-0.828	-2.522	-2.560	-1.939	-1.471	-1.484	-2.074	-1.179	-0.433	0.493
Standard Errors	(1.153)	(1.751)	(2.189)	(2.554)	(2.871)	(3.138)	(3.383)	(3.638)	(3.895)	(4.123)	(4.338)

Figure 1.11: Price Impact of Large Dividends on Exchange Rates

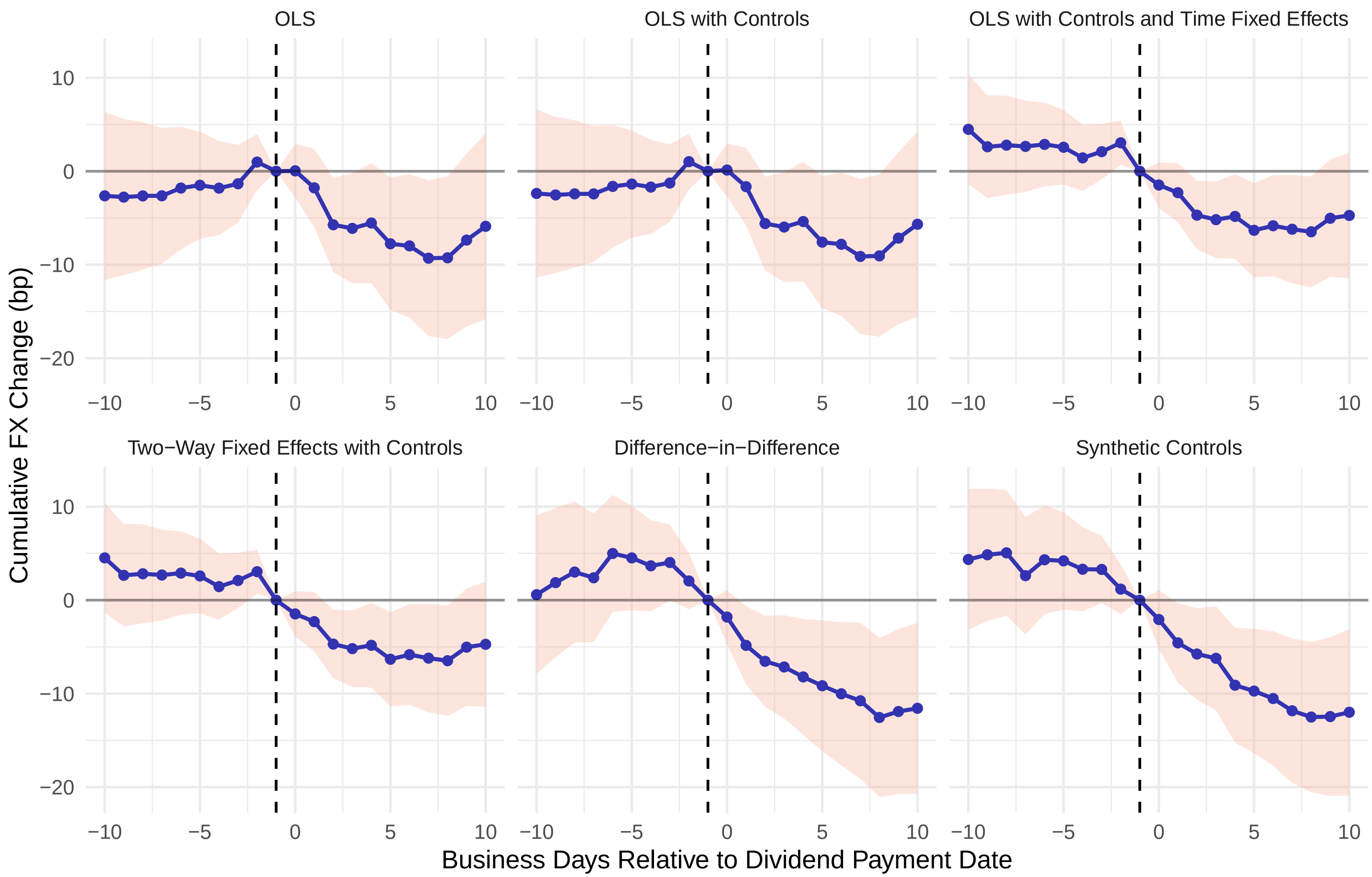


Figure 1.12: Illustration of the Synthetic Control Methodology

This figure illustrates the methodology of estimating the synthetic control, i.e., the best linear combination of control group currencies that best mimics the movement of the treated currency in the estimation window [-70,-11]. The treated unit is the currency that has a top 5% largest dividend payment within a currency-year on the event date. The control group currencies are defined as currencies that do not have top 10% largest dividend payments within a currency-year over the [-10,10] event window. One currency from the control group units is randomly selected to be the placebo. The remaining control group currencies are used for estimation in Eq (1.23).

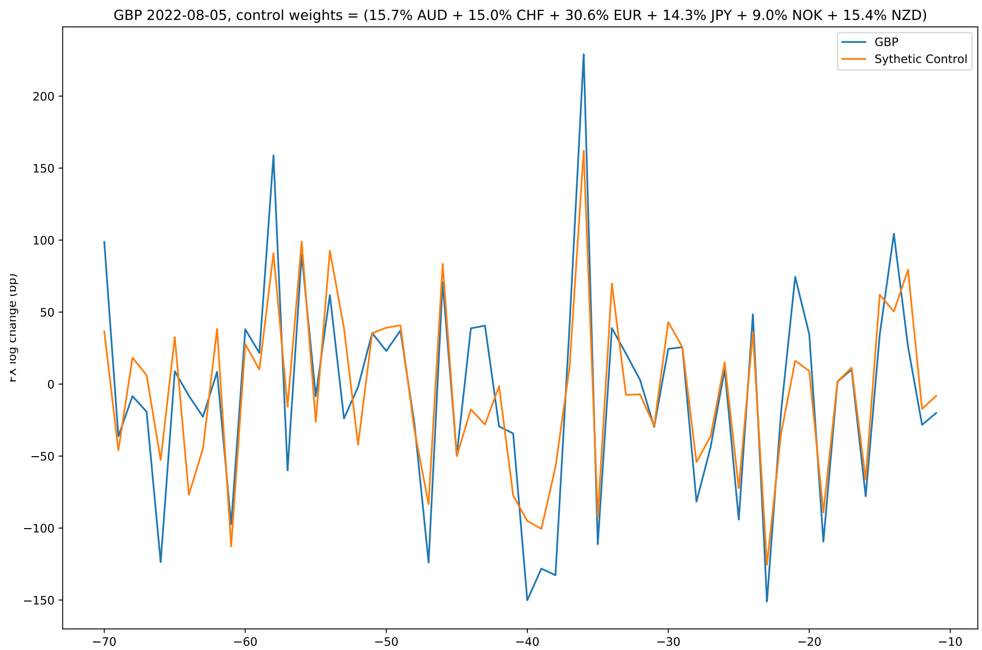


Figure 1.13: Price Impact of Large Dividends on Exchange Rates: Estimates from Difference-in-Difference

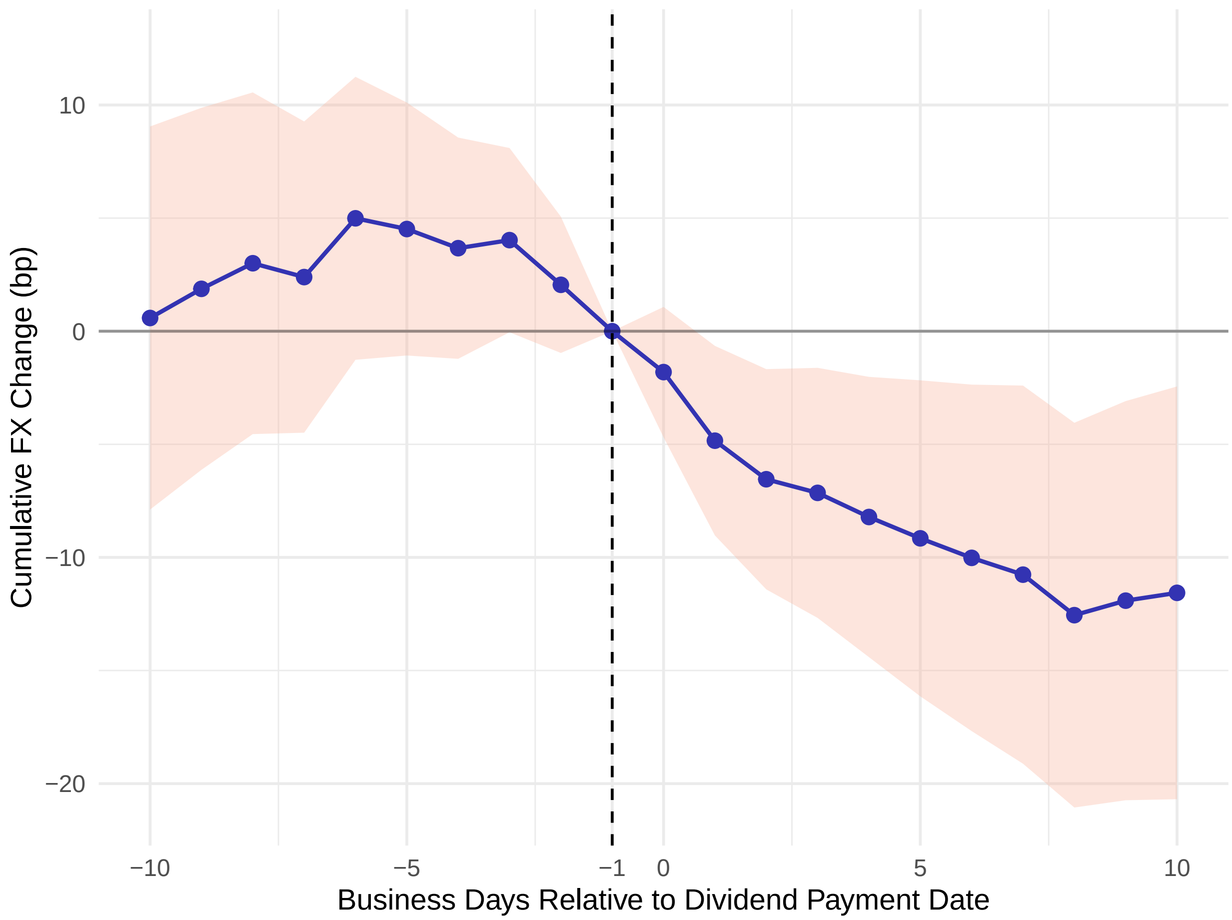


Figure 1.14: Price Impact of Large Dividends on Exchange Rates: Estimates from Synthetic Controls

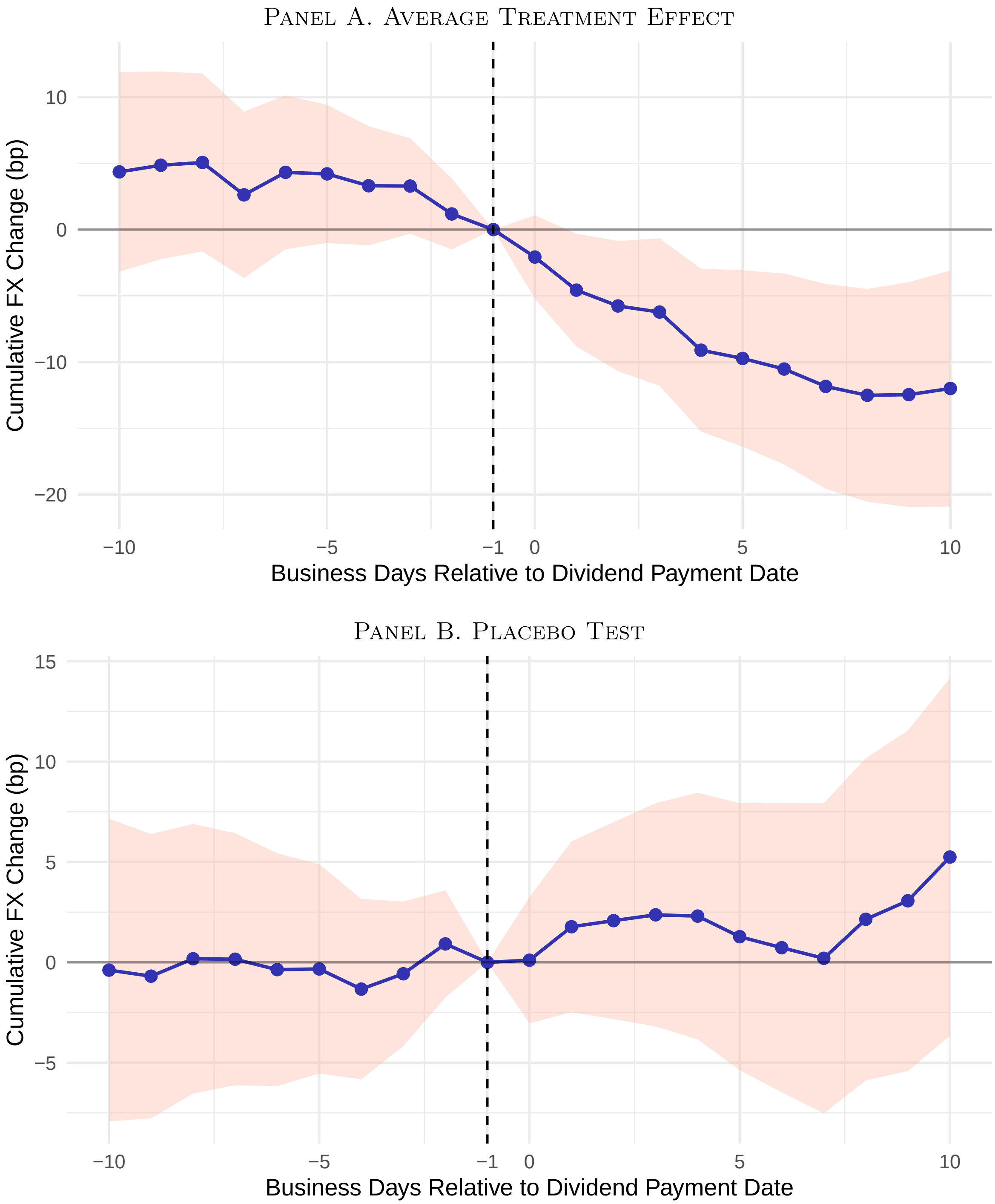


Figure 1.15: Price Impact of Large Dividends on Exchange Rates:
Estimates From DiD by Currency

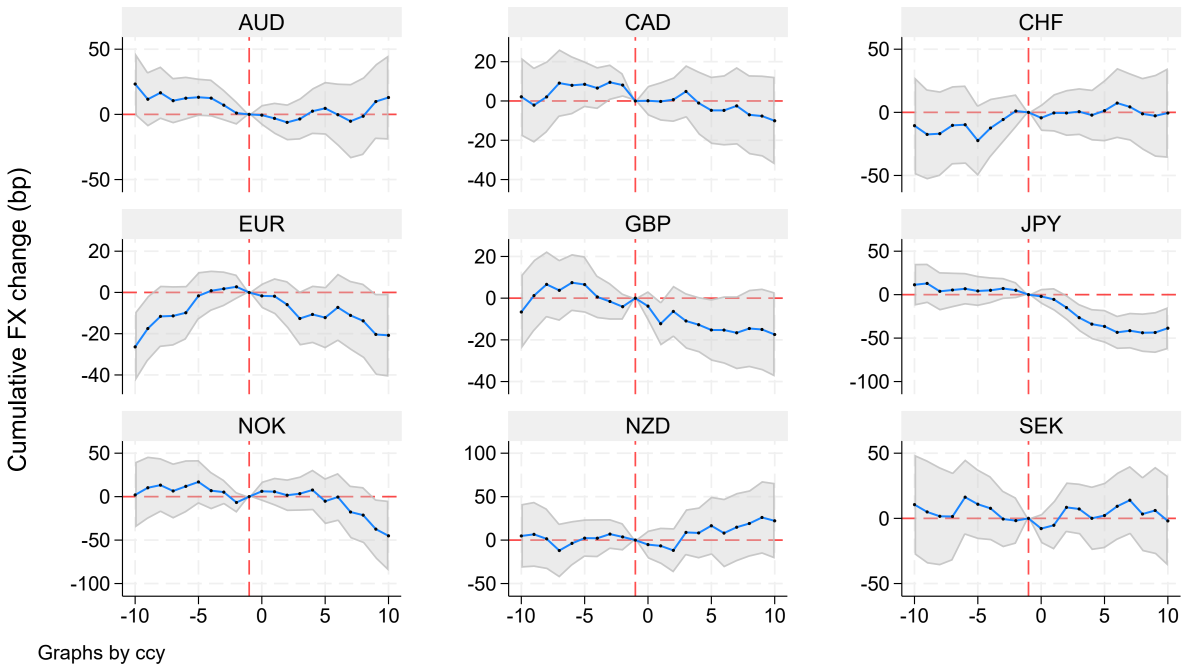


Figure 1.16: Price Impact of Large Dividends on Exchange Rates
Estimates From Synthetic Controls by Currency

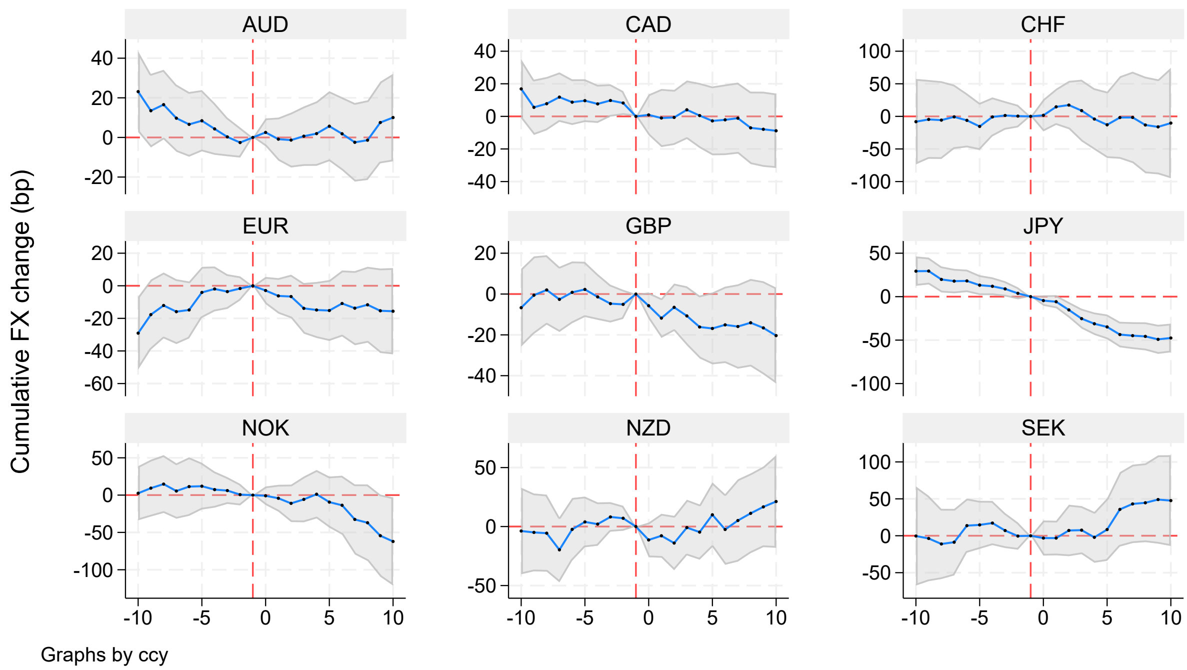
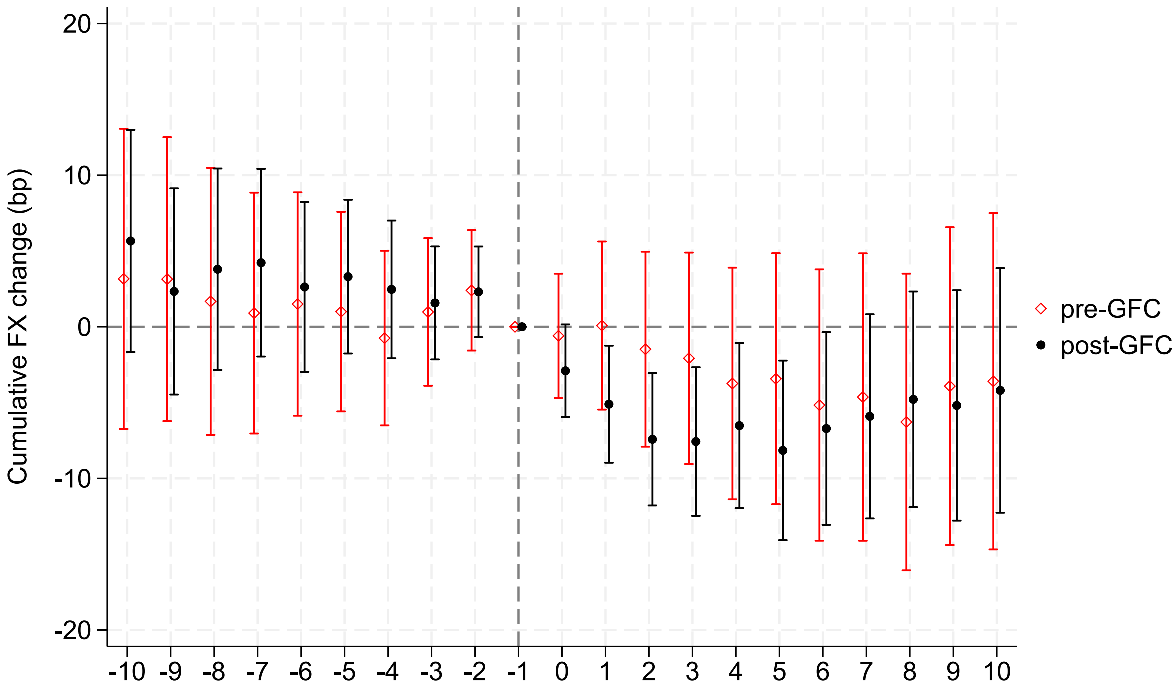


Figure 1.17: Price Impact of Large Dividends on Exchange Rates:
Pre-GFC vs. Post-GFC

This figure compares the coefficients β_h in the baseline regression Eq (1.2) with controls and two way fixed effects, estimated separately before and after the 2007–2008 Global Financial Crisis (GFC). The pre-GFC subsample is from January 2001 to December 2007, and post-GFC subsample is from June 2009 to June 2023, inclusive. The standard errors are two-way clustered at the date level and the currency level.



CHAPTER 2

INVESTORS' DEMAND SYSTEM IN HIGH DIMENSION: A MACHINE LEARNING APPROACH

2.1 Introduction

The investment decision is complex and has become increasingly sophisticated in the age of big data. Each investor has their own style, methodology, or secret in the investment process. They relentlessly collect and analyze new information above and beyond traditional sources to help improve their decision-making. Investors' holdings, as the results of their investment decisions, may help uncover the signals investors track. However, until very recently, the predominant paradigm in asset pricing was to ignore institutional or household holdings data, using only portfolio returns alone or the joint moments of returns and aggregate or individual consumption. Recent initiatives using holdings data, pioneered by Koijen and Yogo (2019), focus on a demand system with a parsimonious factor structure.

In this paper, I develop a high-dimensional characteristics-based demand system that accommodates a large universe of stock characteristics. Importantly, this universe includes price-based stock characteristics, such as momentum and various valuation ratios. Anecdotally, these signals form a significant part of many investors' investment decision-making processes. Therefore, in addition to the omitted bias, excluding these characteristics from demand system estimation may leave a substantial portion of investors' holdings unexplained and absorbed into the demand system's residuals.

Incorporating price-based characteristics may lead to endogeneity issues. Because latent demand is persistent, a price-based characteristic may correlate with latent demand even if the prices involved in its construction are lagged. Suppose there is a positive shock to latent demand at time $t-1$ that raises the stock price at time $t-1$. If the latent demand is persistent, time t latent demand will also be higher, resulting in its correlation with the lagged stock price. This endogeneity could undermine the structural interpretation of coefficients: a non-momentum trader may be "estimated" to be responsive to the momentum signal, overestimating the signal's importance. Therefore, simple OLS or textbook machine learning techniques may lead to biased estimates.

To identify the demand system, I propose an identification strategy based on the intertemporal structure of latent demand to accommodate price-based characteristics, in addition to using the instrumental variable for market equity, as developed in Koijen and Yogo (2019). My identification strategy is structural and motivated by the concept of predeterminedness in the dynamic panel model. Specifically, I assume that the latent demand follows an AR(p) process, and the innovations are mean independent of observables (including prices) before time t. This implies that the stock characteristics are uncorrelated with the future values of innovations in latent demand. In other words, the latent demands from previous periods are sufficient to capture the impact of previously observed variables on the formation of this period's latent demand. Therefore, given previous latent demands in the specification, price-based characteristics can be treated the same as other exogenous characteristics in the demand system, as long as the prices involved are lagged.

To implement the identification strategy in the presence of high-dimensional stock characteristics, I develop a novel machine learning procedure. First, I use the double machine learning (DML) procedure developed in Chernozhukov et al. (2018) to obtain an $\sqrt{N}$-consistent estimate of the price coefficient. Second, I construct another machine learning estimator for the residual demand (demand minus the price term) to achieve the desired statistical properties of the estimand. This approach is especially relevant for estimating investors' demand systems in high-dimensional settings, where the objectives are twofold: 1) $\sqrt{N}$-consistent price coefficient; 2) a combination of stock characteristics, together with price term, robustly fits investors' holdings data, i.e., minimizing the prediction error of the estimated demand system not only in-sample but also out-of-sample.

Standard textbook machine learning (ML) approaches do not deliver a consistent estimate of the price coefficient in demand system estimation because they fail to address price endogeneity and do not incorporate instrumental variables (IV). Alternatively, one may consider an iterated estimator: start with an initial guess of the price coefficient, use the machine learning (ML) technique to estimate the (high-dimensional) controls; then, with the estimated controls, apply IV regression to update the price coefficient estimate, and repeat this process until convergence. Although this intuitive iterated estimation method may provide a consistent estimator, it does not achieve $\sqrt{N}$-consistency. This is due to the heavy bias induced by regularization and overfitting of high-dimensional nuisance parameters (i.e., stock characteristics) on the price coefficient in the demand system. $\sqrt{N}$-consistency of the price coefficient estimate is crucial as it determines the rate of convergence and enables further statistical inference.

To overcome this regularization and overfitting bias, Chernozhukov et al. (2018) proposes a double machine learning (DML) estimator based on Neyman orthogonality conditions and cross-fitting. While DML provides $\sqrt{N}$-consistent estimate for the low-dimensional parameters in the presence of high-dimensional nuisance parameters, its ML estimator for the high-dimensional nuisance parameters serves merely as an intermediate step, instead of providing desired statistical properties for the demand system as a whole. Therefore, with the $\sqrt{N}$-consistent estimate of the price coefficient obtained via DML, I implement an additional ML estimator layer on the demand system. This layer aims to identify a combination of stock characteristics, including the price term, that not only has the desired statistical properties but also robustly fits investors' holdings data. In particular, I choose Lasso based on the sparsity assumption of characteristics in the individual investor's demand system. It is reasonable to assume only a limited number of characteristics influence individual investors' portfolio choice due to inattention (Gabaix 2019), despite in equilibrium the SDF may not be characteristics-sparse (Kozak et al. 2020). Under certain conditions (e.g., Wainwright (2019)

Chapter 7, Zhao and Yu (2006)) , Lasso has the property of variable selection consistency. Using Lasso, we can analyze which characteristics enter into which investors' demand system at different points in time.

In the empirical analysis, I apply the machine learning procedure combined with the identification strategy to estimate the investors' demand system, incorporating a large universe of stock characteristics within the U.S. stock market. The U.S. institutional holdings data are from SEC 13F filings. All institutional investment managers, including foreign institutional investment managers, are required to file Form 13F within 45 days after the end of each calendar quarter, if they use the U.S. mail (or other means or instrumentality of interstate commerce) in the course of their business and that exercise investment discretion over $100 million or more in Section 13(f) securities. The universe of characteristics, consisting of 94 stock-level attributes, is adopted from Green et al. (2017) and constructed using data from CRSP, Compustat, and I/B/E/S.

The estimated demand system shows some stock characteristics, e.g. *LNbe* (log book equity), *divA_be* (dividends to book equity), *mve_ia* (industry-adjusted size), *mom12m* (12-month momentum), *mom36m* (36-month momentum), *cash* (cash holdings), *cashpr* (cash productivity), *sp* (sales to price), *ms* (financial statement score), are almost always selected into demand system of every investor, indicating their strong explanatory power for investors' portfolio holdings. In contrast, *profit* (profitability) and *Gat* (investment), proposed in Fama and French (2015) and used in Koijen and Yogo (2019), only appear in 60% investors' estimated demand. Meanwhile, some characteristics rarely enter investors' demand, indicating their explanatory power for investors' holdings are absorbed by other characteristics. Though most of the top selected characteristics are persistent over time, some characteristics have significant time variation in their presence in investors' demand system, e.g. *beta* (market beta) starts to fall out of favor since the 2000s. On average, the majority of investors have 20-40 characteristics selected in their demand system.

I illustrate the asset pricing relevance of the investors' demand system in a high-dimensional setting through two applications. First, I refine the variance decomposition technique from Koijen and Yogo (2019) to assess the variable importance of each stock characteristic in terms of its impact on equilibrium pricing. Although many characteristics that are frequently present in investors' demand systems, they often exhibit low explanatory power for cross-sectional stock returns. Therefore, a stock characteristic's explanatory power for an investor's holdings does not necessarily translate into its explanatory power in equilibrium asset pricing. For the influential characteristics, we can group them into four categories: 1) price trends, including *mom12m* (12-month momentum) and *indmom* (industry momentum); 2) valuation ratios and fundamental signals, including *LNbe* (log book equity), *bm_ia* (industry-adjusted book to market), *profit* (profitability), *roaq* (return on assets); 3) liquidity variables, including *zerotrade* (number of zero trading days), *baspread* (bidask spread), *turn* (share turnover), *std_turn* (volatily of share turnover); 4) risk measures, including *idiovol*, *beta* (market beta) and *betasq* (beta squared). Among all characteristics in the demand system, *mom12m* (12-month momentum) is the most important characteristic that contributes to cross-sectional stock returns, followed by *LNbe* (log book equity). Interestingly, these influential characteristics are similar to the influential stock-level predictors for monthly stock returns identified by various machine learning techniques in Gu et al. (2020). However, even accounting for such a large universe of observable characteristics documented in the literature, the latent demand's importance in explaining cross-sectional stock returns only reduces moderately, from 81% identified in Koijen and Yogo (2019) to 75.3% in equally-weighted variance decomposition or to 53.4% in the value-weighted case. Therefore, quantitatively, the latent demand is still a puzzle.

Second, I derive a formula for characteristic pricing from the estimated demand system, which sheds light on how changes in stock characteristics relate to changes in equilibrium pricing. A characteristic can significantly influence cross-sectional returns only if two conditions are met: 1) the aggregate demand elasticity with respect to the stock characteristic is large; 2) the aggregate demand with respect to price is inelastic. To see these two effects more transparently, I also derive a first-order approximation for these two effects, which are related to average price coefficient or characteristic loadings, weighted by investors' share of market cap. More interestingly, using the demand system framework, I derive characteristic pricing decomposition into different investors' contribution, in order to answer the question which investors matter for the pricing associated with a stock characteristic. For example, in terms of characteristic pricing associated with *mom12m* (12-month momentum), mutual funds contribute the most in their total effects, followed by investment advisors. The investment advisors' role in pricing *mom12m* is impressive compared to its AUM, given that it only accounts from 5% to 20% of total AUM throughout the sample.

The remainder of the paper is organized as follows. After a brief review of the related literature and comments on my paper's contribution, Section 2.2.1 introduces the log-linear specification of the characteristics-based demand system in high dimension. Section 2.2.2 explains the identification strategy used to identify price and characteristics coefficients. Section 2.2.3 develops the machine learning procedure that fits nicely in the demand system estimation in the high-dimensional setting. Section 2.3.1 details the data used in this paper and relevant transformation. Section 2.3.2 contains other implementation details. Section 2.4.1 summarizes the patterns of demand system estimates from the machine learning estimation procedure. Section 2.4.2 and 2.4.3 illustrate its uses in terms of asset pricing. Finally, Section 2.5 concludes.

2.1.1 Literature Review

My paper is related to several strands of literature. First, this paper is related to the nascent literature on using the demand system approach in asset pricing. Koijen and Yogo (2019) provides a micro-foundation for the characteristics-based demand system under the assumption that returns have a factor structure and that expected returns and factor loadings depend on the assets' own characteristics. Moreover, it proposes an instrumental variable estimator based on cross-sectional variation in the investment universe to address the endogeneity of demand and asset prices. Koijen et al. (2023) uses the demand system approach to study which investors matter for stock valuation. It finds that hedge funds and small active investment advisors are most influential per dollar of assets under management. In contrast, long-term investors, such as pension funds and insurance companies are the least influential. Koijen and Yogo (2024) applies the demand system approach to the international setting. It estimates a demand system for financial assets across short-term debt, long-term debt and equity in 36 countries. It finds that macro variables and policy variables (i.e., short-term rates, debt quantities, and foreign exchange reserves) account for 55 percent of the variation in exchange rates, 57 percent of long-term yields, and 69 percent of stock prices. In all these papers, the demand system specification is parsimonious in characteristics. My paper's contribution to this literature is threefold. First, I propose an identification strategy that admits price-based characteristics. Second, I develop a machine learning procedure that is well-suited for estimating the demand system with a large universe of characteristics. Third, I derive formulae that shed light on characteristic pricing and its decomposition into investors' heterogeneous contributions.

Second, my paper is related to the literature on machine learning in asset pricing studying a high dimension of characteristics. e.g., Green et al. (2017) compiles 94 firm characteristics using Fama-Macbeth regressions and find that 8 to 12 characteristics are significant independent determinants of average returns. This dataset is widely used in the literature (e.g., Gu et al. (2020), Feng et al. (2020)). Freyberger et al. (2020) propose a group lasso procedure to select characteristics and estimate how they affect expected returns non-parametrically. Kozak et al. (2020) constructs a robust stochastic discount factor (SDF) by imposing an economically motivated prior on SDF coefficients that shrinks contributions of low-variance principal components of the candidate characteristics-based factors. Importantly, it finds that characteristics-sparse SDFs like five-factor models cannot adequately summarize the cross-section of expected stock returns. In contrast, PC-sparse SDF, which includes five optimally selected PC-based factors, performs uniformly better both in-sample and out-of-sample. Kelly et al. (2019) proposes Instrumented Principal Component Analysis (IPCA). It allows for latent factors and time-varying loadings by introducing observable characteristics that instrument for the unobservable dynamic loadings. It finds that in the U.S. stock market, five IPCA factors explain the cross-section of average returns well. Moreover, among a large universe of characteristics explored in the literature, only ten are statistically significant in the IPCA specification. Gu et al. (2020) performs a comparative analysis of machine learning methods in predicting returns and demonstrates large economic gains to investors using machine learning forecasts. These papers focus on explaining or predicting returns. My paper enlarges the scope of this literature by focusing on investors' high-dimensional demand system estimation, which in turn has asset pricing implications in equilibrium. To my knowledge, my paper is the first paper that uses machine learning techniques to systematically analyze investors' holdings data.

Finally, this paper is related to recent econometrics literature on inference on low-dimensional parameters in the presence of high-dimensional nuisance parameters. Belloni et al. (2014) proposes the "post-double-selection" method for inference on the effect of a treatment variable in the context of very many controls, under the assumption of approximate sparsity. Belloni et al. (2012) covers the selection of many instruments for IV models with a small number of controls. Chernozhukov et al. (2015) generalizes both papers into the setting with many controls and many instruments. It overcomes the effect of model selection mistake on the parameter of interest by using moment conditions that satisfy Neyman orthogonality condition, i.e., locally insensitive to the value of the high-dimensional nuisance parameters. It also provides an illustrative example in estimating the demand curve in Berry et al. (1995) setting. It finds that by including a large set of control variables, the estimate of the price coefficient is more elastic than the baseline case with only parsimonious controls, indicating omitted variable bias. Beyond lasso-type estimator, Chernozhukov et al. (2018) proposes a double/debiased machine learning estimator, which provides $\sqrt{N}$-consistent estimator for low dimensional parameters of interest in the presence of high-dimensional nuisance parameters. By relying upon Neyman orthogonality condition and cross-fitting, it admits the use of a broad array of modern machine learning techniques, such as lasso, ridge, elastic-net, random forest, boosted trees, deep neural nets, and various hybrids and ensembles of these methods. My paper is the first paper to apply these tools to estimate the demand system in empirical asset pricing.

2.2 Methodology

2.2.1 Model Setup

An (institutional) investor $i \in I$ allocates its wealth $A_{i,t}$ to its investment universe $\mathcal{N}_{i,t}$ and an outside asset $n = 0$. Let $w_{i,t}(n)$ be its portfolio weight on asset $n \in \mathcal{N}_{i,t}$ at date t. Koijen and Yogo (2019) shows that the optimal portfolio simplifies to a characteristics-based demand function that depends on observed characteristics and latent demand (i.e., characteristics unobserved by the econometrician), if 1) returns have a factor structure, and 2) expected returns and factor loadings depend on the assets' own characteristics. Moreover, under a particular coefficient restriction, the relative portfolio weights are exponential-linear in characteristics:

$$\delta_{i,t}(n) := \frac{w_{i,t}(n)}{w_{i,t}(0)} = \exp(\beta_{0,i,t} me_t(n) + \sum_{k=1}^{K-1} \beta_{k,i,t} x_{k,t}(n) + \beta_{K,i,t}) \epsilon_{i,t}(n) \tag{2.1}$$

Combined with the investors' budget constraints implies the following multinomial logit form for portfolio weights:

$$w_{i,t}(n) = \frac{\delta_{i,t}(n)}{1 + \sum_{m \in \mathcal{N}_{i,t}} \delta_{i,t}(m)}$$
$$w_{i,t}(0) = \frac{1}{1 + \sum_{m \in \mathcal{N}_{i,t}} \delta_{i,t}(m)}$$

Taking the logarithm of (2.1) leads to a linear specification:

$$\ln \delta_{i,t}(n) = \beta_{0,i,t} me_t(n) + \sum_{k=1}^{K-1} \beta_{k,i,t} x_{k,t}(n) + \beta_{K,i,t} + \ln \epsilon_{i,t}(n) \tag{2.2}$$

Normalize $\mathrm{E}[\ln \epsilon_{i,t}(n)] = 0$ so that the intercept $\beta_{K,i,t}$ is identifiable. Here, the expectation operator $\mathrm{E}[\cdot]$ is taken over n. Appendix F in Koijen and Yogo (2019) compares the exponential-linear specification (2.1) with the log-linear specification (2.2), and it concludes the difference is small from a value-weighted perspective. For this reason and ease of computation, I will use the log-linear specification (2.2) as my benchmark specification in this paper. More generally, we can assume that the right-hand-side of (2.1) and (2.2) takes the following form:

$$\ln \delta_{i,t}(n) = \beta_{0,i,t} me_t(n) + g(\mathbf{x}_t(n)) + \ln \epsilon_{i,t}(n) \tag{2.3}$$

where $g(\mathbf{x}_t(n)) := \sum_{k=1}^{K-1} \beta_{k,i,t} x_{k,t}(n) + \beta_{K,i,t}$ in (2.2)

Motivated by Fama and French (2015), Koijen and Yogo (2019) includes log book equity, profitability, investment, dividends to book equity, and market beta as characteristics $\{x_{k,t}(n)\}_{k=1}^{5}$, due to concerns about overfitting and collinearity in larger models with more characteristics. Further, Koijen and Yogo (2019) stays away from price-based characteristics because these characteristics could violate their identifying assumption that all characteristics are exogenous to latent demand.

In this paper, I propose an identification strategy based on the assumption of the inter-temporal structure of latent demand, discussed in Section 2.2.2. The identification strategy facilitates the inclusion of price-based characteristics such as momentum and valuation ratios, which are informative to both discretionary and quantitative investors. Additionally, I develop a machine learning procedure in Section 2.2.3 that is well-suited to demand system estimation in high-dimension. This procedure yields a $\sqrt{N}$-consistent estimate of the price coefficient and selects controls that, along with the price term, robustly fit investors' holdings with desired statistical properties.

2.2.2 Identification Strategy

The challenge of identification in estimating (2.2) stems from the endogeneity of stock price in $me_t(n)$ and the endogeneity of price-based stock characteristics like momentum and valuation ratios. The first type of endogeneity, common in demand estimation, biases the estimated demand elasticity with respect to price. To address this issue, Koijen and Yogo (2019) introduces an instrumental variable for market equity, leveraging cross-sectional variations in the investment universe among investors and variations in the sizes of potential investors across assets:

$$\widehat{me}_{i,t}(n) = \log\left(\sum_{j\neq i} A_{j,t}\frac{\mathbb{I}_{\{n\in\mathcal{N}_{j,t}\}}}{1+\sum_{m=1}^{N}\mathbb{I}_{\{m\in\mathcal{N}_{j,t}\}}}\right) \tag{2.4}$$

This instrument variable depends only on other investors' investment universe and the wealth distribution, which are exogenous under the identifying assumptions that wealth distribution across other investors is predetermined and exogenous to current demand shocks. The instrument variable can be interpreted as the counterfactual market equity if other investors were to hold an equal-weighted portfolio within their investment universe. Given a downward-sloping demand curve, a larger exogenous component of demand leads to higher prices that are unrelated to latent demand.

The second endogeneity, i.e., the correlation between price-based characteristics and latent demand, will bias the demand elasticity with respect to characteristics. A price-based characteristic $x_{k,t}(n)$ can be correlated with latent demand $\ln \epsilon_{i,t}(n)$ even when the prices involved in its construct are lagged, due to latent demand persistence. For example, suppose a positive shock to the latent demand $\ln \epsilon_{i,t-1}(n)$ at time $t-1$ leads to a higher stock price $p_{t-1}(n)$. If the latent demand is persistent, time t latent demand $\ln \epsilon_{i,t(n)}$ will also be higher, resulting in its correlation with lagged stock price $p_{t-1}(n)$. Because of this endogeneity, a non-momentum trader can be "estimated" to be responsive to the momentum signal, exaggerating the importance of the latter. If this endogeneity issue is not adequately addressed, it hinders the incorporation of many stock characteristics that many discretionary and quantitative investors closely monitor.

In this paper, I propose an identification strategy to resolve the second type of endogeneity. The identifying assumption is based on the inter-temporal structure of latent demand, i.e., previous periods' latent demands $\{\ln \epsilon_{i,t-l}(n)\}_{l=1}^{p}$ are sufficient to summarize the impact of previous observable variables on the formation of current latent demand. Therefore, innovations in latent demand are uncorrelated with the previous-period observables.

Assumption 1. *The latent demand in* (2.2) *follows AR(p):*

$$\ln \epsilon_{i,t}(n) = \sum_{l=1}^{p} \rho_{l,i,t} \ln \epsilon_{i,t-l}(n) + v_{i,t}(n) \tag{2.5}$$

Further, the innovation $v_{it}(n)$ is mean-independent of observable variables before time t.

This identification assumption is similar to the concept of predeterminedness, where covariates are uncorrelated with future values of structural error. Given that stock characteristics are derived from observable variables, including prices, prior to time t, Assumption 1 implies $E[v_{i,t}(n)|\mathbf{x}_t(n), \{\ln \epsilon_{i,t-l}(n)\}_{l=1}^{p}] = 0$. It should be noted that while the auto-regression parameters $\rho_{l,i,t}$'s in Assumption 1 are constant across stocks n, they may vary across investors i and over time periods t.

Plugging (2.5) into (2.2) or (2.3) leads to the main specification to be estimated:

$$\ln \delta_{i,t}(n) = \beta_{0,i,t} me_t(n) + \underbrace{\sum_{k=1}^{K-1} \beta_{k,i,t} x_{k,t}(n) + \beta_{K,i,t}}_{g(\mathbf{x}_t(n))} + \sum_{l=1}^{p} \rho_{l,i,t} \ln \epsilon_{i,t-l}(n) + v_{i,t}(n) \quad (2.6)$$

with the moment conditions:

$$\mathrm{E}[v_{i,t}(n) | \widehat{me}_i(n), \mathbf{x}_t(n), \{\ln \epsilon_{i,t-l}(n)\}_{l=1}^{p}] = 0 \quad (2.7)$$

The form, represented by $g(\mathbf{x}_t(n))$, admits more general machine learning algorithms, by the machine learning procedure.

2.2.3 Machine Learning Procedure

In this section, I discuss the machine learning procedure for estimating (2.6) with moment conditions (2.7), in the setting with high-dimensional characteristics $\mathbf{x}_{i,t}$. This procedure aims to deliver two key outcomes: 1) a $\sqrt{N}$-consistent estimate of the price coefficient $\beta_{0,i,t}$, and 2) a combination of stock characteristics $\mathbf{x}_{i,t}$ that, alongside the price term, robustly fit investors' holdings data. These two conditions are essential for four reasons: 1) The price coefficient directly relates to demand elasticity, a fundamental parameter in demand estimation. This coefficient determines how an investor's demand responds to changes in stock price. 2) Demand elasticity is also crucial in understanding characteristic pricing in equilibrium. Changes in characteristics cause changes in aggregate demand, which is associated with the price adjustment to achieve a new equilibrium. Demand elasticity with respect to price determines the extent of such price adjustment. 3) Using a broad array of characteristics enhances the demand system's ability to align with investors' holdings data. By minimizing the residuals, we extract investors' demand that can be explained by observable stock characteristics as much as possible. This may help resolve the latent demand puzzle. 4) However, the extensive use of characteristics raises concerns about overfitting. Addressing overfitting requires a systematic model selection strategy, where machine learning techniques such as regularization and cross-validation play critical roles. These techniques ensure a robust fit of the demand system, minimizing prediction errors both in-sample and out-of-sample.

As the procedure is general and goes beyond the linear case, I will present the framework in the partially linear IV model, using the notations in Chernozhukov et al. (2018). For an investor i at time t, we can rewrite (2.6) :

$$\begin{aligned} &Y = D\theta_0 + g_0(X) + U, \quad E[U|X, Z] = 0 \\ &Z = m_0(X) + V, \quad E[V|X] = 0 \end{aligned} \tag{2.8}$$

where the correspondence with our previous notations is:

$$\begin{aligned} &Y := \ln \delta_{i,t}(n), D := me_t(n), \theta_0 = \beta_{0,i,t}, U = v_{i,t} \\ &g(X) = \sum_{k=1}^{K-1} \beta_{k,i,t} x_{k,t}(n) + \beta_{K,i,t} + \sum_{l=1}^{p} \rho_{l,i,t} \ln \epsilon_{i,t-l}(n) \\ &Z := \widehat{me}_{i,t}(n), X := (\mathbf{x}_{i,t}, \{\ln \epsilon_{i,t-l}(n)\}_{l=1}^{p}) \end{aligned}$$

Intuitively, one might consider using iterative estimators: starting with an initial guess of the price coefficient, apply a machine learning (ML) technique to estimate controls that fit the demand minus the price term. Following this, use IV regression with these controls to update the price coefficient estimate, repeating the process until convergence. However, Chernozhukov et al. (2018) points out this iterative estimator does not deliver a $\sqrt{N}$-consistent estimate of the price coefficient, since the machine learning estimator of $g_0(X)$ introduces significant bias in the estimator of θ_0, due to regularization bias and overfitting. Consequently, this bias causes the intuitive iterated estimator $\hat{\theta}_0$ to fail in achieving $\sqrt{N}$-consistency.

Therefore, I adapt the double/debiased machine learning (DML) in Chernozhukov et al. (2018) to the demand system estimation in high-dimensional settings, which is $\sqrt{N}$-consistent and admits the use of a broad array of machine learning methods, including random forests, lasso, ridge, deep neural nets, boosted trees, and various hybrids and ensembles of these methods. The procedure has two critical ingredients: 1) Neyman orthogonal scores that are locally insensitive to the value of the high-dimensional nuisance parameters; 2) cross-fitting, an efficient form of data-splitting, to remove bias induced by overfitting. Note that the use of cross-fitting is to control overfitting for the use of a much broader collection of ML methods for estimating the nuisance functions. For lasso-type methods under sparsity or approximate sparsity conditions, the cross-fitting is not needed for $\sqrt{N}$-consistency (Chernozhukov et al. 2015).

I re-state the definition of double/debiased machine learning estimator, i.e. Definition 3.2 in Chernozhukov et al. (2015), adapted to the partially linear IV model. For a sample $W = (Y, D, X, Z)$ and functions $\eta = (\ell(X), m(X), r(X))$, define the Robinson-style Neyman orthogonal score for partially linear IV model (2.8) as [1]:

$$\psi(W; \theta, \eta) = (Y - \ell(X) - \theta(D - r(X)))(Z - m(X)) \tag{2.9}$$

At the true value $\eta_0 = (\ell_0(X), m_0(X), r_0(X))$, where $\ell_0 = \mathrm{E}[Y|X]$, $r_0(X) = \mathrm{E}[D|X]$, and $m_0(X) = \mathrm{E}[Z|X]$, this function satisfies 1) moment condition: $\mathrm{E}[\psi(W; \theta_0, \eta_0)] = 0$; 2) orthogonality condition: $\partial_\eta \mathrm{E}[\psi(W; \theta_0, \eta_0)][\eta - \eta_0] = 0$.

Definition 2.2.1 (DML). *(a) Take a K-fold random partition of observations indices $[N] = \{1, \ldots, N\}$ such that the size of each fold I_k is $n = N/K$. For each $k \in \{1, \ldots, K\}$, define $I_k^c = \{1, \ldots, N\} \setminus I_k$. (b) For each $k \in [K]$, construct an ML estimator $\hat{\eta}_{0,k} = \hat{\eta}_0((W_i)_{i \in I_k^c})$*

1. Chernozhukov et al. (2018) provides another choice of the Neyman score in this context: $\psi(W; \theta, \eta) = (Y - \theta D - g(X))(Z - m(X))$. However, the Robinson-style Neyman orthogonal score is preferred, as the nuisance parameters involved are conditional mean functions, which can be directly estimated by ML methods.

of η_0 using data on the sub-sample I_k^c. (c) Construct the estimator $\tilde{\theta}_0$ as the solution to: [2]

$$\frac{1}{K}\sum_{k=1}^{K} \mathrm{E}_{n,k}[\psi(W;\tilde{\theta}_0,\hat{\eta}_{0,k}] = 0 \tag{2.10}$$

where ψ is the Neyman orthogonal score defined in (2.9)*, and* $\mathrm{E}_{n,k}[\psi(W)] = \frac{1}{n}\sum_{i\in I_k}\psi(W_i)$ *is the empirical expectation over the k-th fold of the observations.*

The DML estimator provides us with $\tilde{\theta}_0$, a $\sqrt{N}$-consistent estimate for θ_0. However, the ML estimators $\{\hat{\eta}_{0,k}\}_{k=1}^{K}$ of $\eta_0 = (\mathrm{E}[Y|X], \mathrm{E}[Z|X], \mathrm{E}[D|X])$ serve only as intermediate steps in estimating θ_0 and do not directly correspond to $g_0(X)$, which, together with $D\theta_0$, aims to minimize the error term. To achieve the aforementioned dual goals, I propose the following machine learning procedure:

Definition 2.2.2 (Estimation Procedure)**.** *(a) Construct the double machine learning (DML) estimator $\hat{\theta}_0$ using the procedure in Definition 2.2.1. (b) Construct the ML estimator $\hat{g}_0(X)$ for* $\mathrm{E}[Y - D\theta|X]|_{\theta=\hat{\theta}_0}$.

The above estimation procedure allows for the use of general machine learning techniques, including lasso, ridge, elastic-net, random forests, boosted trees, deep neural nets, and various hybrids and ensembles of these methods. Moreover, one can use a different choice of machine learning techniques for $\ell(X), m(X), r(X)$ and $g(X)$, further enhancing the procedure's flexibility. Chernozhukov et al. (2018) outlines some considerations for selecting machine learning techniques: 1) Approximate sparsity suggests the use of forward selection, lasso, post-lasso, L2-boosting, or some other sparsity-based technique; 2) Well-approximability by trees indicates the suitability of regression trees and random forest; 3) Well-approximability by sparse neural and deep neural nets calls for the use of ℓ_1-penalized neural and deep neural

2. An alternative choice of DML estimator $\hat{\theta}_0$ is the average of $\check{\theta}_{0,k}$, where $\check{\theta}_{0,k}$ is the solution to $\mathrm{E}_{n,k}[\psi(W;\check{\theta}_{0,k},\hat{\eta}_{0,k})] = 0$. However, Chernozhukov et al. (2018) recommends the use of the DML estimator stated in the main text, as it is better behaved due to the more stable behavior of pooled empirical Jacobians.

networks; 4) Well-approximability by at least one of the models above justifies the use of an ensemble/aggregated method.

In the demand system estimation, the (approximate) sparsity assumption of characteristics in the individual investor's demand system is a reasonable assumption. Due to investors' inattention (Gabaix 2019), only a limited number of characteristics influence individual investors' portfolio choice, although in equilibrium, the SDF may not be characteristics-sparse (Kozak et al. 2020). In addition, under certain conditions (e.g., Wainwright (2019) Chapter 7, Zhao and Yu (2006)), Lasso has the property of variable selection consistency, i.e., the support set selected by Lasso coincides with the true support set. Using Lasso, we can analyze which characteristics enter into which investors' demand system at different points in time. For these reasons, I choose the Lasso-type machine learning estimator in the following empirical analysis.

2.3 Empirical Analysis

2.3.1 Data

The underlying data in my paper combines the data in Koijen and Yogo (2019) and Green et al. (2017). I use the merged CRSP-Compustat database and the Thomson Reuters Institutional Holdings (13F) Database for replication of Koijen and Yogo (2019), and I use the merged CRSP-Compustat and the I/B/E/S database for replication of Green et al. (2017). The data sample covers the U.S. ordinary common shares (i.e., share codes 10, 11, 12, and 18) that trade on the New York Stock Exchange, the American Stock Exchange, and Nasdaq (i.e., exchange codes 1, 2, and 3) from 1980Q1 to 2017Q4.

The data on institutional common stock holdings are from the Thomson Reuters Institutional Holdings (13F) Database, as reported on Form 13F filed with the SEC. It contains quarterly stock holdings of institutions since 1980. All institutional investment managers, including foreign institutional investment managers, must file Form 13F within 45 days after the end of each calendar quarter, if they use the United States mail (or other means or instrumentality of interstate commerce) in the course of their business and exercise investment discretion over $100 million or more in Section 13(f) securities. Collectively, these institutions manage 35 percent of the U.S. stock market in 1980-1984 to 68 percent in 2015-2017. Note that Form 13F reports only long positions but not short positions. Moreover, short positions in a security are not subtracted from long positions. Option positions may be reported, but the 13F data only contains long positions and does not have the details of strikes and expiration dates. We also do not know the cash and bond positions of institutions because these assets are not part of 13(f) securities.

Following Koijen and Yogo (2019), I group the institutions into six types: 1) banks, 2) insurance companies, 3) investment advisors, 4) mutual funds, 5) pension funds, and 6) other 13F institutions. An investment advisor is a registered company under Securities and Exchange Commission Form ADV, including many hedge funds. Investment advisors that are mutual funds form a separate group. Other 13F institutions include endowments, foundations, and non-financial corporations. The household sector is defined as the difference between shares outstanding and the sum of shares held by 13F institutions. This difference represents direct household holdings and smaller institutions that are not required to file Form 13F. Assets Under Management (AUM) is the sum of dollar holdings for each institution. Figure 2.1 shows the evolution of AUM by investor types and their proportion of the total market capitalization. While household (as defined above) is still the biggest sector, its proportion has steadily decreased from 68% in 1980 to 33% in 2017, while investment advisors and mutual funds increase from 5% to 20-30% at the end of the sample.

The universe of stock characteristics includes 94 firm characteristics in Green et al. (2017). These characteristics are also used in Gu et al. (2020), Feng et al. (2020). For log book equity, dividends to book equity, profitability, investment, and market beta, I use the construct in Koijen and Yogo (2019) to facilitate comparison.[3] For exposition purposes, I refer to these five characteristics as KY characteristics and the remaining 89 characteristics as GHZ characteristics.

To mitigate the influence of outliers, I perform rank normalization on the GHZ characteristics, mapping them onto the [-0.5, +0.5] interval, where the lowest percentile corresponds to -0.5 and the highest percentile corresponds to +0.5. This transformation is a common practice in the literature, e.g. Asness et al. (2019), Freyberger et al. (2020), Kelly et al. (2019), Kozak et al. (2020), Gu et al. (2020). Under this normalization, the magnitude of $\beta_{k,i,t}$ can be interpreted as the effect on asset n's log relative weight, if its characteristic $x_{k,t}(n)$ moves from the bottom to the top in the cross-section. Note that I do *not* perform the rank normalization on KY characteristics, allowing their estimated coefficients to be directly comparable to the results in Koijen and Yogo (2019). For any missing values in the stock characteristics, I fill these gaps with the cross-sectional median for each characteristic at each date, see Section 2.3.1 for further details.

To avoid the forward-looking bias, Green et al. (2017) lags monthly characteristics by 1 month, quarterly characteristics by 4 months, and annual characteristics by 6 months in the constructs of stock characteristics. However, as the granularity of time in Assumption 1 for the empirical analysis of 13F data is one quarter, all prices involved in the constructs of stock characteristics $x_{i,t}(n)$ should be at least one quarter ago. Therefore, I make the following changes in GHZ characteristics. First, I further lag *mom12m* (11-month cumulative returns ending one month before month end) so that it represents cumulative returns from t-14m to t-3m. Second, I drop *maxret* (maximum daily return from returns during calendar month $t-1$), *retvol* (standard deviation of daily returns from month $t-1$), *mom1m* (1-month cumulative return), drop *mom6m* (5-month cumulative returns ending one month before month end), *chmom* (change in 6m momentum, i.e. cumulative returns from months t-6

3. Accordingly, I exclude the characteristics *bm*, *dy*, *operprof*, *agr*, *beta* from Green et al. (2017), respectively.

to t-1 minus months t-12 to t-7) from the universe of characteristics used, since lagging these variables by 2 more months results in variables that have low correlation with original constructs, and may not capture the original motivation underlying these variables. Third, I drop *mve* (natural log of market capitalization at end of month $t-1$), as this is highly correlated with $me_t(n)$. Fourth, I drop *dolvol* (natural log of trading volume times price per share from month $t-2$) and related *std_ dolvol* (monthly standard deviation of daily dollar trading volume), as there is strong multi-collinearity between stock price, *dolvol* and *turn* (average monthly trading volume for most recent 3 months scaled by number of shares outstanding in current month). Lastly, I drop *ill* (average of daily absolute return/dollar volume) as its presence causes lasso estimates to be unstable in the implementation. Note that other illiquidity measures are still present in the universe, e.g., *baspread* (monthly average of daily bid-ask spread divided by average of daily spread). Table 2.1 lists the remaining 85 characteristics. The details of each variable's construct are in the Appendix of Green et al. (2017).

2.3.2 Implementation Details

For the auto-regression order of the inter-temporal structure of latent demand in Assumption 1, I use $p=1$ for the main results in this paper. The results are quantitatively similar for $p=2,3,4$. The investors' demand systems are estimated quarter by quarter. To do so, I iterate the estimation of (2.6) over time. At time t, given the previous period estimated latent demand $\widehat{\ln \epsilon}_{i,t-1}(n)$, use the machine learning procedure outlined in Definition 2.2.2 for $\sqrt{N}$-consistent $\beta_{0,i,t}$ and the high-dimensional controls. During this procedure, $\widehat{\ln \epsilon}_{i,t-1}(n)$ are partialled-out so that the AR(1) structure of the latent demand is preserved and the estimate of its loading $\rho_{1,i,t}$ won't be biased or penalized due to regularization. The latent demand at time t is estimated to be $\widehat{\ln \epsilon}_{i,t}(n) = \hat{\rho}_{1,i,t}\widehat{\ln \epsilon}_{i,t-1}(n) + \hat{v}_{i,t}(n)$, which is then used for the estimation in the next iteration at time $t+1$. As the data sample starts from 1980Q1,

$\forall i, n$, I set $\ln \epsilon_{i,1979Q4}(n) = 0$ as initialization.

I use post-double-selection lasso (*pdslasso*) for step (a) and lasso for step (b) in the machine learning estimation procedure outlined in Definition 2.2.2. For the double machine learning in part (a), Chernozhukov et al. (2015) shows that under the assumption of approximate sparsity, we do not need to use cross-fitting to obtain a $\sqrt{N}$-consistent estimator. Therefore, I set $K = 1$ for the number of observation partitions. For the Lasso in part (b), I use standard cross-validation with 3 folds.[4] The use of lasso is for two reasons:[5] 1) Koijen and Yogo (2019) shows that under a particular form of coefficient restriction, the demand system is exponential-linear in characteristics. 2) the (approximate) sparsity assumption of characteristics in the individual investor's demand system is a reasonable assumption, due to investors' inattention (Gabaix 2019).

To facilitate direct comparison with Koijen and Yogo (2019), especially the results on variance decomposition, I follow the procedure of pooled estimates for small institutions. In other words, for institutions with more than 1,000 strictly positive holdings, I estimate the demand system at the individual institution level. For institutions with fewer than 1,000 holdings, I group them by type and quantiles of AUM conditional on type, and estimate the demand system at the group level. Therefore, all institutions within a group have the same coefficients in their demand system estimation. It is worth mentioning that the machine learning procedure described in Definition 2.2.2 is applicable at the individual level for investors with a limited number of holdings.

4. One can also use rLasso which uses a data-driven rate-optimal penalty.

5. That being said, as shown in as shown in Gu et al. (2020), non-linearity and interaction between characteristics do matter in the settings of return prediction. The importance of non-linearity in demand system is left for future research.

2.4 Main Results

2.4.1 Estimated Demand System

Table 2.2 presents the price and characteristic coefficients estimated using the machine learning procedure described in Definition 2.2.2, equally weighted for all investors and for each investor type. Columns *freq* refer to how often a particular characteristic is selected into the estimated demand system. Some stock characteristics are almost always selected for all investors, including *LNbe* (log book equity), *divA_be* (dividends to book equity), *mve_ia* (industry-adjusted size), *mom12m* (12-month momentum), *mom36m* (36-month momentum), *cash* (cash holdings), *cashpr* (cash productivity), *sp* (sales to price), and *ms* (financial statement score). Notably, *profit* (profitability) and *Gat* (investment) used in Koijen and Yogo (2019) only enters around 60% percent of investors' estimated demand system.

Figure 2.2 illustrates the frequency with which stock characteristics are selected into the investors' demand system over time, weighted by investors' AUM. A darker color indicates more frequent selection. We can draw three conclusions. First, the most frequently selected characteristics align with the equal-weighted results previously discussed. Second, though there are some variations over time, most top selected characteristics are highly persistent. One notable exception is *beta* (market beta), which has declined in popularity since the 2000s. Third, the infrequent selection of many characteristics corroborates the sparsity assumption inherent in lasso-type estimators.

Figure 2.3 illustrates the distribution of the number of characteristics selected in investors' demand systems. For most investors, the number of characteristics typically ranges from 20 to 40 over time. This indicates that incorporating a broader set of characteristics beyond the KY attributes more effectively matches investors' holdings data.

2.4.2 Variance Decomposition

In this section, I refine the variance decomposition in Koijen and Yogo (2019), which decomposes the log stock returns into contribution by changes in 1) shares outstanding, 2) stock characteristics, 3) dividend yield, 4) assets under management, 5) coefficients on characteristics, and 6) latent demand:

$$\mathbf{r}_{t+1} = \mathbf{p}_{t+1} - \mathbf{p}_t + \mathbf{v}_{t+1} \tag{2.11}$$

$$\begin{aligned}\mathrm{Var}(\mathbf{r}_{t+1}) = {} & \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\mathbf{s}), \mathbf{r}_{t+1}) + \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\mathbf{x}), \mathbf{r}_{t+1}) + \mathrm{Cov}(\Delta\mathbf{v}_{t+1}, \mathbf{r}_{t+1}) \\ & + \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\mathbf{A}), \mathbf{r}_{t+1}) + \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\beta), \mathbf{r}_{t+1}) + \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\epsilon), \mathbf{r}_{t+1})\end{aligned} \tag{2.12}$$

where

$$\begin{aligned}
\Delta\mathbf{p}_{t+1}(\mathbf{s}) &:= \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_t, \mathbf{A}_t, \beta_t, \epsilon_t) - \mathbf{g}(\mathbf{s}_t, \mathbf{x}_t, \mathbf{A}_t, \beta_t, \epsilon_t) \\
\Delta\mathbf{p}_{t+1}(\mathbf{x}) &:= \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_{t+1}, \mathbf{A}_t, \beta_t, \epsilon_t) - \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_t, \mathbf{A}_t, \beta_t, \epsilon_t) \\
\Delta\mathbf{p}_{t+1}(\mathbf{A}) &:= \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_{t+1}, \mathbf{A}_{t+1}, \beta_t, \epsilon_t) - \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_{t+1}, \mathbf{A}_t, \beta_t, \epsilon_t) \\
\Delta\mathbf{p}_{t+1}(\beta) &:= \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_{t+1}, \mathbf{A}_{t+1}, \beta_{t+1}, \epsilon_t) - \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_{t+1}, \mathbf{A}_{t+1}, \beta_t, \epsilon_t) \\
\Delta\mathbf{p}_{t+1}(\epsilon) &:= \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_{t+1}, \mathbf{A}_{t+1}, \beta_{t+1}, \epsilon_{t+1}) - \mathbf{g}(\mathbf{s}_{t+1}, \mathbf{x}_{t+1}, \mathbf{A}_{t+1}, \beta_{t+1}, \epsilon_t)
\end{aligned}$$

Here, $\mathbf{g}(\mathbf{s}, \mathbf{x}, \mathbf{A}, \beta, \epsilon)$ is the vector of counter-factual prices defined by market clearing conditions, given number of shares $\mathbf{s}$, stock characteristics $\mathbf{x}$, investors wealth $\mathbf{A}$, coefficients on characteristics β and latent demand ϵ:

$$\mathbf{q}_t := \log(\sum_{i=1}^{I} A_{i,t}\mathbf{w}_{i,t}) - \mathbf{p}_t = \mathbf{s}_t \tag{2.13}$$

Koijen and Yogo (2019) finds that among all these effects, latent demand changes are the most important, explaining 81 percent of the cross-sectional variance of stock returns. They conclude that stock returns are mostly explained by demand shocks that are unrelated to changes in observed characteristics, i.e., the latent demand puzzle.

Does including a broader universe of characteristics help reduce the importance of latent demand? Can additional stock characteristics described in the literature help explain the cross-sectional variance of stock returns? If so, which stock characteristics are important? With the estimated high-dimensional demand system at hand, I proceed to answer these questions through the refined variance decomposition. Specifically, I rewrite the price change $\Delta\mathbf{p}_{t+1}(\mathbf{x})$ induced by changes in stock characteristics as a series of price changes induced by changing one stock characteristic at a time:

$$\Delta\mathbf{p}_{t+1}(\mathbf{x}) = \sum_{k=1}^{K}\left[\Delta\mathbf{p}_{t+1}(\mathbf{x}_{t\to t+1}^{(k)}) - \Delta\mathbf{p}_{t+1}(\mathbf{x}_{t\to t+1}^{(k-1)})\right] \tag{2.14}$$

where $\mathbf{x}_{t\to t+1}^{(k)}$ denotes changing the first k characteristics while keeping the last $N-k$ characteristics the same:

$$\mathbf{x}_{t\to t+1}^{(k)} = \begin{bmatrix} x_{1,t+1}(1) & \dots & x_{k,t+1}(1) & x_{k+1,t}(1) & \dots & x_{K,t}(1) \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{1,t+1}(N) & \dots & x_{k,t+1}(N) & x_{k+1,t}(N) & \dots & x_{K,t}(N) \end{bmatrix}$$

The variance of log stock returns (2.12) can be further decomposed into:

$$\begin{aligned} \mathrm{Var}(\mathbf{r}_{t+1}) = {} & \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\mathbf{s}), \mathbf{r}_{t+1}) + \sum_{k=1}^{K}\mathrm{Cov}(\left[\Delta\mathbf{p}_{t+1}(\mathbf{x}_{t\to t+1}^{(k)}) - \Delta\mathbf{p}_{t+1}(\mathbf{x}_{t\to t+1}^{(k-1)})\right], \mathbf{r}_{t+1}) \\ & + \mathrm{Cov}(\Delta\mathbf{v}_{t+1}, \mathbf{r}_{t+1}) \\ & + \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\mathbf{A}), \mathbf{r}_{t+1}) + \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\beta), \mathbf{r}_{t+1}) + \mathrm{Cov}(\Delta\mathbf{p}_{t+1}(\epsilon), \mathbf{r}_{t+1}) \end{aligned} \tag{2.15}$$

where $\mathrm{Var}(\cdot)$ and $\mathrm{Cov}(\cdot,\cdot)$ are performed element-wise for vectors in (2.12) and (2.15).

Table 2.3 presents the refined variance decomposition of cross-sectional annual stock returns pooled together, based on the estimated demand system. The contribution of changes in shares outstanding, assets under management, coefficients on characteristics are similar to Koijen and Yogo (2019). However, most stock characteristics contribute very little in magnitude to the variance decomposition. The influential stock characteristics can be categorized into four groups: 1) price trends, including *mom12m* (12-month momentum) and *indmom* (industry momentum); 2) valuation ratios and fundamental signals, including *LNbe* (log book equity), *divA_be* (dividends to book equity), *profit* (profitability), *roaq* (return on assets); 3) liquidity variables, including *zerotrade* (number of zero trading days), *baspread* (bidask spread), *turn* (share turnover), *std_turn* (volatily of share turnover); 4) risk measures, including *idiovol*, *beta* (market beta) and *betasq* (beta squared). Among all characteristics in the demand system, *mom12m* (12-month momentum) is the most important characteristic that contributes to cross-sectional stock returns, followed by *LNbe* (log book equity). Interestingly, these influential characteristics are similar to the influential predictors for monthly stock returns identified by various machine learning techniques in Gu et al. (2020), despite the primary focus is to explain investors' demand instead of return prediction. Therefore, the variance decomposition approach, as proposed in Koijen and Yogo (2019) and further refined in this paper, serves as a tool for assessing each characteristic's importance in equilibrium asset pricing and pinpointing key stock characteristics that are pivotal.

Despite incorporating a comprehensive set of characteristics and employing a machine learning approach to estimate the demand system, the latent demand still contributes to the majority of cross-sectional variance of stock returns (75.3% for equally weighted and 53.4% for value-weighted, compared with 81% in Koijen and Yogo (2019)). Quantitatively, the latent demand is still a big puzzle! What exactly constitutes latent demand? How does it relate to investors' belief? What factors influence shifts in latent demand? Delving into these questions and exploring latent demand beyond observable stock characteristics is an important topic for future research.

2.4.3 Characteristic Pricing

In this section, I derive a formula of characteristics pricing from the estimated demand system. In addition, I decompose the heterogeneous contributions of investors and analyze which investors matter for equilibrium pricing of a stock characteristic.

Take the derivative of (2.13) with respect to the vector of the k-th characteristic for all stocks $\mathbf{x}_{k,t}$, we can calculate the equilibrium stock price change associated with a change in percentile of the stock characteristic, holding other characteristics constant:

$$\frac{\partial \mathbf{p}_t}{\partial \mathbf{x}'_{k,t}} = \left[-\frac{\partial \mathbf{q}_t}{\partial \mathbf{p}'_t}\right]^{-1} \frac{\partial \mathbf{q}_t}{\partial \mathbf{x}'_{k,t}} \tag{2.16}$$

The first term on the right-hand side, $\left[-\frac{\partial \mathbf{q}_t}{\partial \mathbf{p}'_t}\right]^{-1}$, is the elasticity of aggregate demand with respect to price:

$$-\frac{\partial \mathbf{q}_t}{\partial \mathbf{p}'_t} = \mathbf{I} - \sum_{i=1}^{I} \beta_{0,i,t} A_{i,t} \mathbf{H}_t^{-1} \mathbf{G}_{i,t} \tag{2.17}$$

The second term is the elasticity of aggregate demand with respect to a stock characteristic, which can be written similarly as:

$$\frac{\partial \mathbf{q}_t}{\partial \mathbf{x}'_t} = \sum_{i=1}^{I} \beta_{k,i,t} A_{i,t} \mathbf{H}_t^{-1} \mathbf{G}_{i,t} \tag{2.18}$$

where $\mathbf{G}_{i,t} = \text{diag}(\mathbf{w}_{i,t}) - \mathbf{w}_{i,t}\mathbf{w}'_{i,t}$ is related to investor i's portfolio allocation, and $\mathbf{H}_t = \sum_{i=1}^{I} A_{i,t}\text{diag}(\mathbf{w}_{i,t})$ is a diagonal matrix whose elements are market capitalization for each stock.

From equation (2.16), a stock characteristic influences cross-sectional returns if a change in the characteristic leads to a shift in aggregate demand for some stocks, and via own-price elasticity and/or cross-price elasticity, relates to price changes in equilibrium. More specifically, if a stock characteristic $\mathbf{x}_k$ does not enter investors' demand system (i.e. $\forall i \in I, \beta_{k,i,t} \approx 0$), then this characteristic is not priced in equilibrium. Even a stock characteristic enters into investors' demand system, its price impact may be minimal, if the weighted sum $\sum_{i=1}^{I} \beta_{k,i,t} A_{i,t} \mathbf{H}_t^{-1} \mathbf{G}_{i,t} \approx 0$. Furthermore, if a characteristic only impacts the demand of stocks whose own-price elasticity and cross-price elasticity are large, then this characteristic does not significantly affect the cross-sectional stock prices. On the contrary, a characteristic that enters only a few investors' demand system could still matter for cross-sectional returns, as long as (2.16) is significant.

Moreover, we can use (2.16) to decompose the pricing associated with a stock characteristic into contribution by different investors or investor types. Partition the investor indices as $I = \bigcup_{g=0}^{6} I_g$, which refers to 0) households, 1) banks, 2) insurance companies, 3) investment advisors, 4) mutual funds, 5) pension funds, and 6) other institutions, respectively. By grouping the summation in the second term in (2.16) by investor types, we have:

$$\frac{\partial \mathbf{p}_t}{\partial \mathbf{x}'_{k,t}} = \sum_{g=0}^{6} \left[-\frac{\partial \mathbf{q}_t}{\partial \mathbf{p}'_t} \right]^{-1} \left(\sum_{i \in I_g} \beta_{k,i,t} A_{i,t} \mathbf{H}_t^{-1} \mathbf{G}_{i,t} \right) \tag{2.19}$$

Each of the term $\left[-\frac{\partial \mathbf{q}_t}{\partial \mathbf{p}'_t} \right]^{-1} \left(\sum_{i \in I_g} \beta_{k,i,t} A_{i,t} \mathbf{H}_t^{-1} \mathbf{G}_{i,t} \right)$ represents an investor group's contribution to the pricing associated with stock characteristic k at time t. To gain further insight into (2.19), I use the first-order approximation of $G_{i,t} \approx \text{diag}(\mathbf{w}_{i,t})$. (2.19) can then be simplified to a diagonal matrix, with n-th diagonal element being:

$$\frac{\partial p_t(n)}{\partial x_{k,t}(n)} = \sum_{g=0}^{6} \left[\left(1 - \sum_{i=1}^{I} \beta_{0,i,t} \frac{A_{i,t} w_{i,t}(n)}{ME_t(n)} \right)^{-1} \left(\sum_{i \in I_g} \beta_{k,i,t} \frac{A_{i,t} w_{i,t}(n)}{ME_t(n)} \right) \right] \tag{2.20}$$

where $ME_t(n) = \sum_{i=1}^{I} A_{i,t} w_{i,t}(n)$ is the market capitalization for stock n. Intuitively, an investor's contribution to the stock price change associated with a change in characteristics is proportional to its characteristic loading, scaled by its share of the stock's market capitalization.

Figure 2.4 illustrates such decomposition for the cross-sectional price changes associated with a change in *mom12m* (12-month momentum). I average $\frac{\partial p_t(n)}{\partial x_{k,t}(n)}$ using (2.19) across stocks n within a year and then average across years for each investor type. The figure suggests mutual funds contribute the most to the characteristic pricing of stock momentum, followed by investment advisors. The investment advisors' role in pricing *mom12m* is substantial relative to their AUM, which comprises only 5% to 20% of the total AUM over the sample period. In contrast, despite households managing the largest share of AUM throughout the sample, their contribution to the characteristic pricing associated with *mom12m* is moderate.

2.5 Conclusion

Estimating investors' demand that accommodates a large universe of stock characteristics is of both academic and practical value. It is becoming increasingly relevant in the age of big data and AI/ML. In this paper, I develop a machine learning procedure to estimate the demand system in high dimension. My contribution is threefold. First, I propose an identification strategy based on the assumption of the inter-temporal structure of latent demand, in order to resolve the endogeneity of price-based stock characteristics. This opens the door to incorporate many closely monitored stock characteristics into the investors' demand system estimation, e.g., momentum and valuation ratios. Second, I develop a machine learning estimation procedure that is well-suited for estimating investors' demand system with a large universe of characteristics. This procedure delivers a $\sqrt{N}$-consistent estimator of price coefficient and a combination of stock characteristics that robustly fits the holdings data and has the desired statistical properties. Third, I explore the asset pricing implications of the high-dimensional demand system. Refining the variance decomposition technique, I identify four groups of the most influential stock characteristics that contribute to the cross-sectional stock returns. Among all characteristics, 12-month momentum is the most important. Moreover, I decompose the characteristic pricing associated with 12-month momentum into contribution by different investor groups. I find that while mutual funds contribute the most, investment advisors play an impressive role, given their proportion of AUM. In contrast, households play a much moderate role in this regard even with the largest AUM throughout the sample.

The latent demand remains a puzzle quantitatively, despite incorporating a large universe of characteristics using machine learning techniques. Overshadowing observable stock characteristics, the latent demand still explains the majority of cross-sectional stock returns. An important topic for future research is to understand the nature of this latent demand.

Figure 2.1: Evolution of Asseet Under Management by Investor Type

The institutions are grouped into six types: 1) banks, 2) insurance companies, 3) investment advisors, 4) mutual funds, 5) pension funds, and 6) other 13F institutions. The household sector is defined as the difference between shares outstanding and the sum of shares held by 13F institutions. An investment advisor is a registered company under Securities and Exchange Commission Form ADV, including many hedge funds. Investment advisors that are mutual funds are separated. Other 13F institutions includes endowments, foundations, and non-financial corporations. Assets under management (AUM) is the sum of dollar holdings for each institution. The quarterly sample period is from 1980Q1 to 2017Q4.

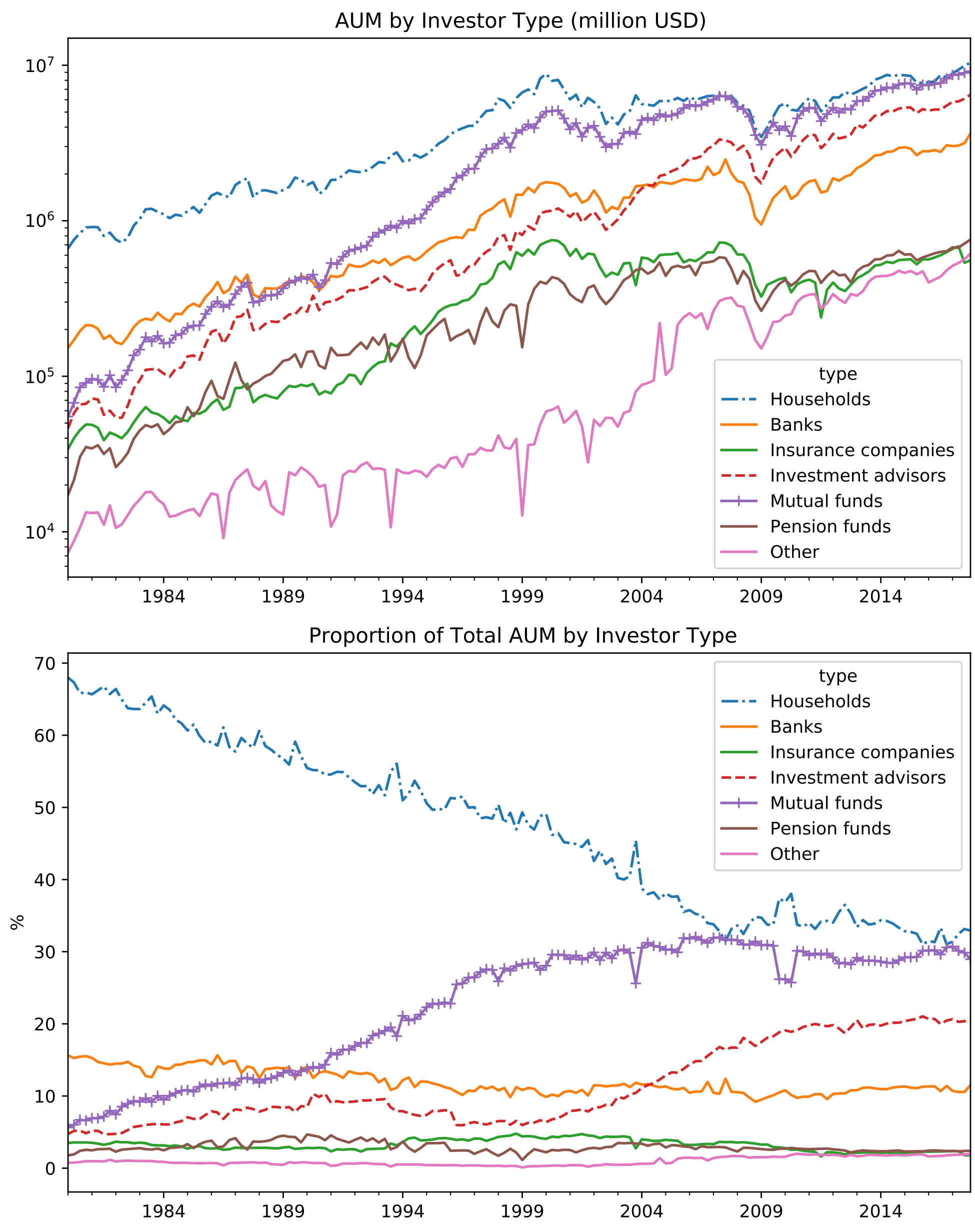


Figure 2.2: Frequency of Characteristic Occurrence in Investors' Demand System

This figure illustrates the frequency with which a stock characteristic is selected in the investor's demand system for each date. I estimate the investor's demand system using the machine learning procedure outlined in Definition 2.2.2 with Lasso, applying it period by period. The estimates are then averaged across investor types, weighted by each investor's Assets Under Management (AUM) at each date. A darker color indicates that a stock characteristic is selected more frequently. The quarterly sample period is from 1980Q1 to 2017Q4.

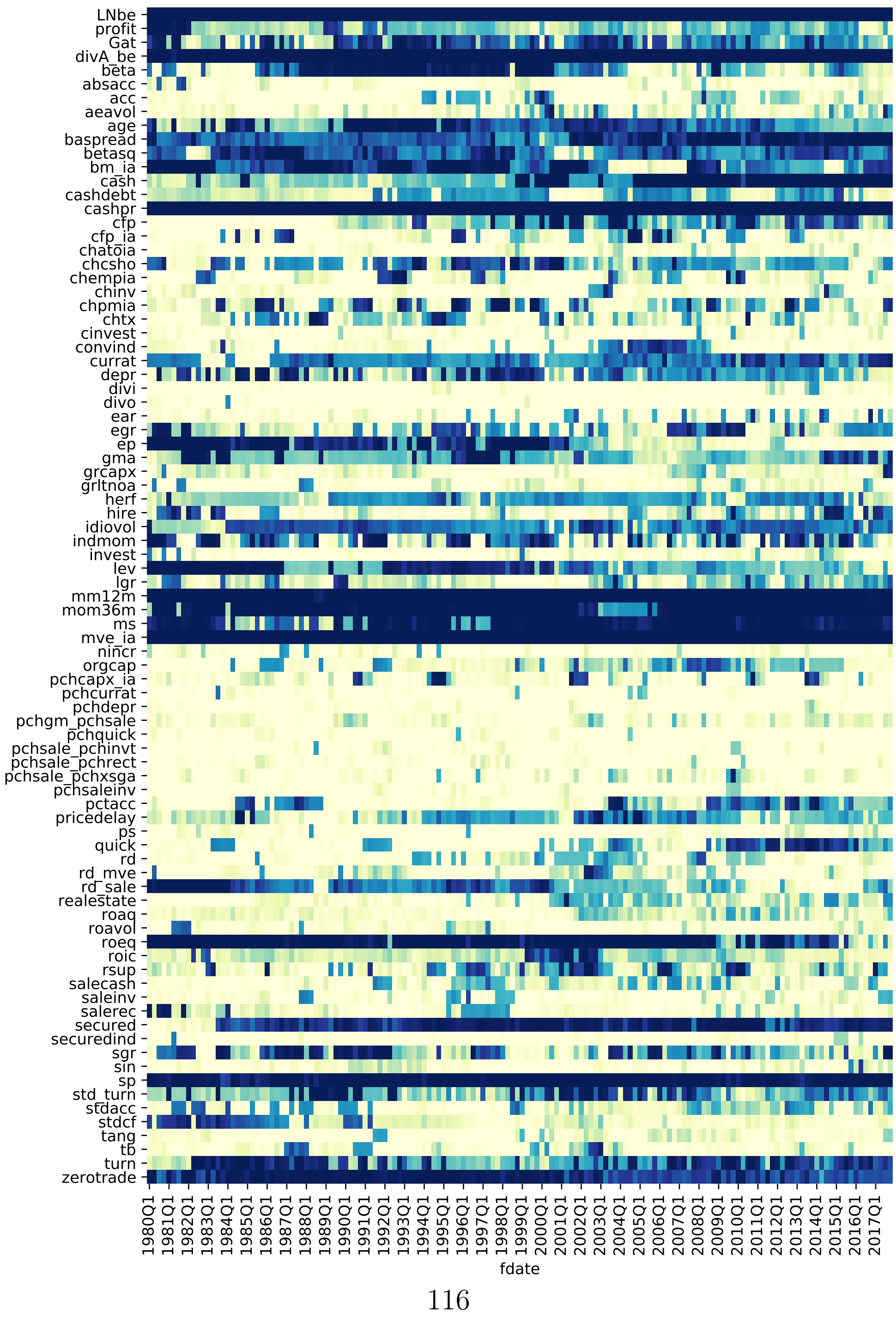


Figure 2.3: Number of Characteristics Selected in Investors' Demand System

This figure displays the distribution of the number of characteristics selected in investors' demand system estimated using the machine learning procedure outlined in Definition 2.2.2. I estimate the investor's demand system for each period separately, then aggregate these estimates over time, by each type of investor. The quarterly sample period is from 1980Q1 to 2017Q4.

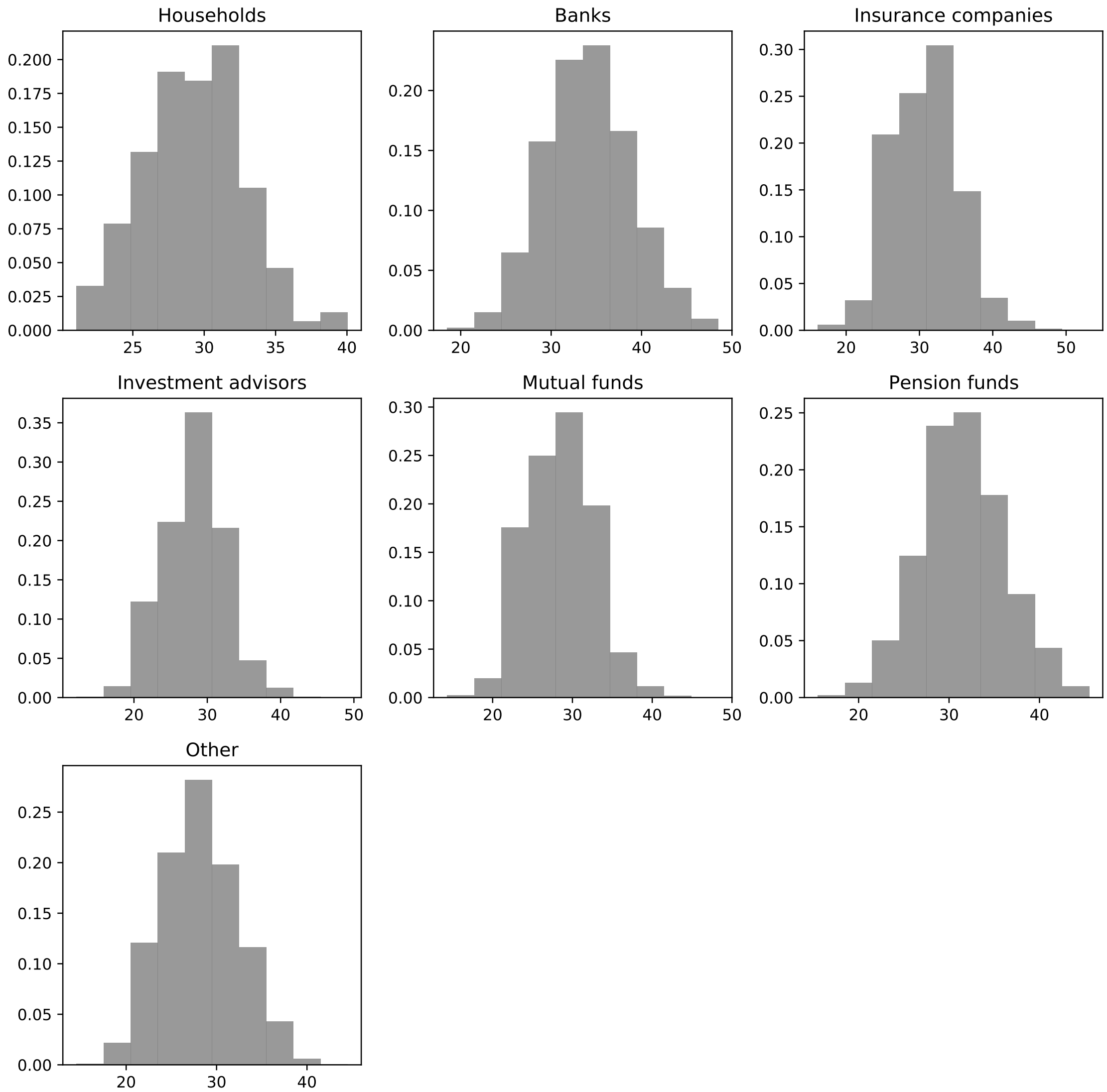


Figure 2.4: Contribution to Characteristic Pricing of 12-Month Momentum by Investor Type

This figure illustrates the contributions of different investor types to the pricing of the stock characteristic *mom12m* (12-month momentum). I average each investor group's contribution to $\frac{\partial p_t(n)}{\partial x_{k,t}(n)}$ using (2.19) across stocks n within a year, and then across years for each investor type. The quarterly sample period is from 1980Q1 to 2017Q4.

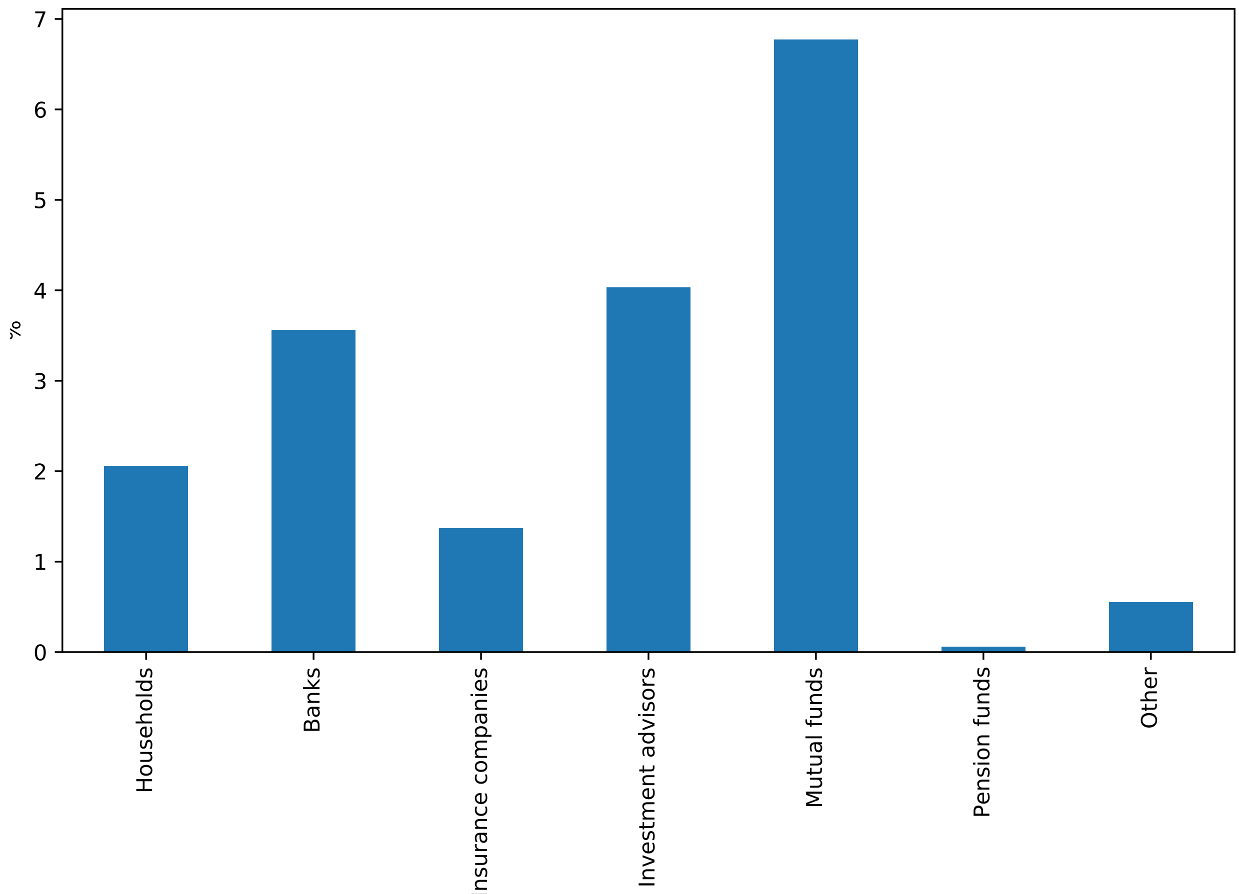


Table 2.1: Details of Stock Characteristics

The table lists the characteristics used in the empirical analysis of the U.S. stock market. *LNbe*, *profit*, *Gat*, *divA_be*, *beta* are the same constructs as in Koijen and Yogo (2019). Other characteristics are adapted from Green et al. (2017) and normalized via rank transformation onto [-0.5, 0.5]. The missing values of each stock characteristic are filled with the cross-sectional median at each date. See Section 2.3.1 for further details.

No.	Characteristic	Description
1	*LNbe*	Log book equity
2	*profit*	Profitability
3	*Gat*	Investment
4	*divA_be*	Dividends to book equity
5	*beta*	Market beta
6	*absacc*	Absolute accruals
7	*acc*	Working capital accruals
8	*aeavol*	Abnormal earnings announcement volume
9	*age*	# years since first Compustat coverage
10	*baspread*	Bid-ask spread
11	*betasq*	Beta squared
12	*bm_ia*	Industry-adjusted book to market
13	*cash*	Cash holdings
14	*cashdebt*	Cash flow to debt
15	*cashpr*	Cash productivity
16	*cfp*	Cash flow to price ratio
17	*cfp_ia*	Industry-adjusted cash flow to price ratio
18	*chatoia*	Industry-adjusted change in asset turnover
19	*chcsho*	Change in shares outstanding
20	*chempia*	Industry-adjusted change in employees
21	*chinv*	Change in inventory
22	*chpmia*	Industry-adjusted change in profit margin
23	*chtx*	Change in tax expense
24	*cinvest*	Corporate investment
25	*convind*	Convertible debt indicator
26	*currat*	Current ratio
27	*depr*	Depreciation / PP&E
28	*divi*	Dividend initiation
29	*divo*	Dividend omission
30	*ear*	Earnings announcement return
31	*egr*	Growth in common shareholder equity
32	*ep*	Earnings to price
33	*gma*	Gross profitability
34	*grcapx*	Growth in capital expenditures
35	*grltnoa*	Growth in long term net operating assets
36	*herf*	Industry sales concentration
37	*hire*	Employee growth rate
38	*idiovol*	Idiosyncratic return volatility
39	*indmom*	Industry momentum
40	*invest*	Capital expenditures and inventory
41	*lev*	Leverage
42	*lgr*	Growth in long-term debt
43	*mom12m*	12-month momentum
44	*mom36m*	36-month momentum
45	*ms*	Financial statement score

(Continued on the next page)

Table 2.1 (Continued): Details of Stock Characteristics

No.	Characteristic	Description
46	*mve_ia*	Industry-adjusted size
47	*nincr*	Number of earnings increases
48	*orgcap*	Organizational capital
49	*pchcapx_ia*	Industry adjusted % change in capital expenditures
50	*pchcurrat*	% change in current ratio
51	*pchdepr*	% change in depreciation
52	*pchgm_pchsale*	% change in gross margin - % change in sales
53	*pchquick*	% change in quick ratio
54	*pchsale_pchinvt*	% change in sales - % change in inventory
55	*pchsale_pchrect*	% change in sales - % change in A/R
56	*pchsale_pchxsga*	% change in sales - % change in SG&A
57	*pchsaleinv*	% change sales-to-inventory
58	*pctacc*	Percent accruals
59	*pricedelay*	Price delay
60	*ps*	Financial statements score
61	*quick*	Quick ratio
62	*rd*	R&D increase
63	*rd_mve*	R&D to market capitalization
64	*rd_sale*	R&D to sales
65	*realestate*	Real estate holdings
66	*roaq*	Return on assets
67	*roavol*	Earnings volatility
68	*roeq*	Return on equity
69	*roic*	Return on invested capital
70	*rsup*	Revenue surprise
71	*salecash*	Sales to cash
72	*saleinv*	Sales to inventory
73	*salerec*	Sales to receivables
74	*secured*	Secured debt
75	*securedind*	Secured debt indicator
76	*sgr*	Sales growth
77	*sin*	Sin stocks
78	*sp*	Sales to price
79	*std_turn*	Volatility of liquidity (share turnover)
80	*stdacc*	Accrual volatility
81	*stdcf*	Cash flow volatility
82	*tang*	Debt capacity/firm tangibility
83	*tb*	Tax income to book income
84	*turn*	Share turnover
85	*zerotrade*	Zero trading days

(Continued from the previous page)

Table 2.2: Estimate for Price and Characteristic Coefficients

This table shows the summary statistics for price coefficient and characteristic coefficients estimated using the procedure in Definition 2.2.2, for all investors and each investor type. The coefficients are equally weighted. Columns *freq* refer to how often a particular characteristic is selected in the estimated demand systems over time. The quarterly sample period is from 1980Q1 to 2017Q4.

	All			Households			Banks			Insurance companies		
	freq	mean	sd	freq	mean	sd	freq	mean	sd	freq	mean	sd
LNme	100%	0.25	0.26	100%	0.58	0.18	100%	0.17	0.43	100%	0.32	0.31
LNbe	100%	0.02	0.18	100%	0.20	0.12	100%	0.25	0.29	100%	0.10	0.20
profit	66%	0.18	0.25	14%	0.03	0.17	85%	0.29	0.34	81%	0.18	0.28
Gat	58%	0.07	0.19	57%	-0.03	0.07	55%	0.06	0.21	62%	0.04	0.18
divA_be	100%	-0.57	2.59	100%	4.63	2.46	100%	3.43	3.18	100%	0.38	2.19
beta	39%	-0.02	0.11	63%	0.01	0.04	56%	0.00	0.13	53%	0.00	0.12
absacc	11%	-0.01	0.14	2%	0.03	0.02	17%	0.01	0.15	14%	0.00	0.09
acc	10%	0.01	0.19	20%	-0.02	0.08	8%	0.02	0.25	6%	-0.02	0.11
aeavol	23%	0.01	0.12	7%	-0.05	0.06	28%	0.00	0.11	25%	0.02	0.09
age	55%	-0.05	0.18	82%	0.24	0.24	69%	-0.15	0.29	70%	-0.06	0.22
baspread	75%	-0.28	0.33	99%	-0.02	0.14	60%	-0.18	0.31	53%	-0.23	0.29
betasq	38%	-0.05	0.23	92%	-0.11	0.16	54%	-0.06	0.26	48%	-0.04	0.23
bm_ia	43%	-0.01	0.14	86%	-0.09	0.09	50%	-0.11	0.19	55%	-0.06	0.15
cash	97%	0.03	0.17	44%	0.19	0.19	97%	0.08	0.18	96%	0.04	0.15
cashdebt	22%	0.18	0.23	45%	-0.22	0.20	41%	0.17	0.26	25%	0.05	0.26
cashpr	100%	0.00	0.19	100%	0.04	0.10	100%	0.10	0.22	100%	0.05	0.21
cfp	57%	-0.01	0.21	26%	-0.05	0.10	58%	-0.11	0.23	59%	-0.06	0.20
cfp_ia	23%	0.04	0.11	15%	0.00	0.05	31%	0.07	0.13	27%	0.03	0.11
chatoia	8%	-0.01	0.10	1%	0.01	0.00	10%	-0.03	0.12	9%	-0.02	0.08
chcsho	58%	0.01	0.11	41%	0.01	0.04	51%	-0.02	0.12	49%	0.01	0.10
chempia	19%	0.02	0.11	13%	0.00	0.02	25%	0.02	0.11	23%	0.01	0.11
chinv	15%	0.01	0.10	3%	0.08	0.10	16%	0.03	0.10	13%	0.00	0.11
chpmia	39%	0.00	0.10	33%	0.01	0.04	46%	0.00	0.10	43%	0.01	0.09
chtx	29%	0.02	0.09	19%	0.02	0.04	32%	0.03	0.09	33%	0.03	0.10
cinvest	10%	0.01	0.08	1%	0.04	0.09	14%	0.03	0.08	8%	0.01	0.07
convind	13%	0.00	0.07	19%	-0.26	0.26	16%	0.01	0.06	6%	0.03	0.05
currat	27%	0.03	0.20	91%	-0.21	0.20	14%	0.02	0.18	14%	0.00	0.16
depr	38%	0.06	0.14	63%	-0.07	0.12	70%	0.14	0.16	54%	0.09	0.13
divi	7%	0.00	0.16	0%	.	.	6%	-0.02	0.25	5%	-0.04	0.13
divo	1%	-0.06	0.42	1%	-0.02	.	2%	-0.07	0.57	1%	-0.07	0.08
ear	15%	0.01	0.09	10%	0.01	0.03	16%	0.01	0.08	15%	0.01	0.08
egr	41%	0.05	0.13	34%	-0.03	0.06	45%	0.06	0.15	46%	0.02	0.12
ep	37%	0.04	0.24	54%	0.01	0.05	64%	-0.09	0.29	67%	0.00	0.21
gma	74%	0.05	0.21	26%	-0.07	0.14	92%	0.19	0.23	88%	-0.01	0.20
grcapx	18%	0.03	0.12	2%	-0.04	0.05	24%	0.04	0.11	22%	0.05	0.13
grltnoa	20%	-0.01	0.12	4%	0.02	0.01	18%	-0.03	0.13	17%	-0.02	0.10
herf	20%	0.11	0.16	64%	-0.27	0.18	53%	0.21	0.17	45%	0.09	0.12
hire	30%	0.05	0.12	20%	0.01	0.07	27%	0.05	0.13	26%	0.06	0.13
idiovol	36%	-0.37	0.41	86%	0.25	0.30	62%	-0.68	0.48	45%	-0.43	0.38
indmom	43%	0.03	0.11	56%	0.08	0.08	46%	0.06	0.13	46%	0.04	0.11
invest	9%	0.02	0.14	3%	0.00	0.03	8%	-0.01	0.16	10%	0.01	0.12
lev	40%	-0.01	0.30	78%	-0.04	0.13	71%	-0.01	0.36	70%	0.05	0.29
lgr	45%	0.04	0.12	12%	0.00	0.03	44%	0.07	0.13	48%	0.03	0.11
mm12m	100%	0.17	0.24	100%	0.04	0.11	98%	0.27	0.29	99%	0.19	0.29
mom36m	98%	0.13	0.19	91%	0.01	0.10	97%	0.23	0.24	100%	0.16	0.21
ms	95%	0.07	0.14	87%	-0.02	0.09	94%	0.19	0.14	89%	0.07	0.12
mve_ia	100%	-0.18	0.22	100%	0.19	0.20	100%	-0.23	0.28	100%	-0.14	0.24
nincr	8%	0.00	0.07	3%	0.00	0.01	13%	-0.01	0.08	11%	-0.01	0.07
orgcap	16%	-0.04	0.22	31%	0.03	0.12	13%	0.01	0.30	17%	0.03	0.21
pchcapx_ia	20%	-0.01	0.10	11%	0.01	0.07	26%	-0.03	0.11	25%	0.00	0.09
pchcurrat	3%	0.01	0.14	3%	-0.01	0.13	6%	0.01	0.14	3%	-0.01	0.11
pchdepr	5%	-0.01	0.11	0%	.	.	6%	-0.02	0.12	2%	0.03	0.08
pchgm_pchsale	15%	0.01	0.10	0%	.	.	22%	0.01	0.10	16%	0.03	0.10
pchquick	2%	0.00	0.13	1%	0.12	0.18	6%	-0.03	0.13	3%	-0.02	0.07
pchsale_pchinvt	5%	0.02	0.28	2%	0.03	0.05	6%	0.05	0.22	5%	-0.01	0.17
pchsale_pchrect	7%	0.00	0.10	1%	0.05	.	14%	0.02	0.11	9%	0.04	0.09
pchsale_pchxsga	13%	0.00	0.11	3%	0.04	0.07	12%	0.00	0.12	11%	0.01	0.09
pchsaleinv	4%	0.00	0.30	2%	0.04	0.04	5%	0.02	0.24	4%	0.06	0.16
pctacc	32%	0.00	0.14	32%	0.18	0.16	28%	0.10	0.18	22%	0.03	0.14
pricedelay	30%	0.01	0.11	54%	0.02	0.05	38%	0.04	0.13	33%	0.02	0.11
ps	7%	0.01	0.11	1%	0.00	0.00	8%	-0.01	0.10	5%	-0.01	0.07
quick	27%	0.03	0.20	33%	-0.18	0.29	13%	-0.01	0.20	15%	0.00	0.17
rd	12%	0.00	0.08	25%	0.02	0.04	13%	0.01	0.08	9%	-0.02	0.10
rd_mve	18%	-0.12	0.20	9%	0.09	0.11	20%	-0.25	0.25	20%	-0.13	0.19
rd_sale	19%	0.03	0.20	78%	0.15	0.17	28%	0.13	0.20	32%	0.03	0.16
realestate	38%	-0.01	0.15	2%	-0.01	0.02	31%	0.04	0.19	37%	0.02	0.18
roaq	37%	0.13	0.25	0%	.	.	29%	0.08	0.28	23%	0.11	0.27
roavol	9%	-0.05	0.16	4%	-0.02	0.01	15%	-0.07	0.17	15%	-0.01	0.13
roeq	80%	0.00	0.19	89%	0.01	0.07	89%	0.07	0.19	91%	0.03	0.17
roic	54%	0.13	0.20	12%	-0.03	0.06	62%	0.13	0.24	47%	0.14	0.23
rsup	35%	0.05	0.13	29%	0.02	0.05	32%	0.03	0.12	36%	0.03	0.10
salecash	31%	-0.05	0.22	8%	-0.16	0.30	34%	-0.15	0.22	31%	-0.09	0.21
saleinv	12%	0.01	0.11	11%	0.09	0.08	13%	0.02	0.12	10%	-0.01	0.10
salerec	9%	-0.07	0.16	14%	-0.01	0.03	24%	-0.18	0.17	19%	-0.08	0.11
secured	78%	0.04	0.19	88%	-0.01	0.09	78%	-0.01	0.22	73%	0.02	0.15
securedind	3%	0.01	0.17	3%	-0.06	0.08	6%	0.00	0.08	2%	-0.01	0.07
sgr	47%	0.05	0.14	43%	0.05	0.05	52%	0.07	0.16	57%	0.08	0.15
sin	17%	0.04	0.17	1%	0.19	0.06	29%	0.11	0.18	15%	-0.01	0.19
sp	95%	0.01	0.24	100%	-0.31	0.20	92%	-0.15	0.30	94%	-0.08	0.27
std_turn	68%	0.02	0.22	50%	-0.01	0.12	87%	-0.04	0.21	82%	-0.03	0.20
stdacc	33%	-0.04	0.19	22%	0.00	0.06	32%	-0.10	0.30	27%	-0.07	0.20
stdcf	23%	-0.16	0.25	22%	-0.01	0.04	63%	-0.29	0.28	31%	-0.10	0.19
tang	11%	0.03	0.13	4%	0.10	0.14	9%	0.09	0.15	10%	0.02	0.12
tb	14%	0.04	0.11	10%	0.01	0.05	23%	0.06	0.10	16%	0.02	0.09
turn	70%	0.02	0.41	56%	-0.64	0.66	84%	-0.18	0.43	76%	-0.01	0.38
zerotrade	61%	-0.08	0.47	100%	0.04	0.34	73%	0.09	0.50	69%	-0.03	0.42
cons	100%	-2.59	1.46	100%	-3.96	0.55	100%	-3.90	1.68	100%	-3.56	1.71
Observations	259093			152			26760			8709		

(Continued on the next page)

Table 2.2 (Continued): Estimate for Price and Characteristic Coefficients

	Investment advisors			Mutual funds			Pension funds			Other		
	freq	mean	sd	freq	mean	sd	freq	mean	sd	freq	mean	sd
LNme	100%	0.23	0.19	100%	0.36	0.26	100%	0.51	0.32	100%	0.22	0.25
LNbe	100%	-0.01	0.12	100%	-0.05	0.18	100%	0.08	0.17	100%	0.05	0.15
profit	63%	0.16	0.21	68%	0.18	0.26	81%	0.10	0.27	49%	0.14	0.24
Gat	58%	0.08	0.18	57%	0.08	0.20	67%	-0.01	0.15	57%	0.07	0.18
divA_be	100%	-1.01	1.82	100%	-2.07	2.48	100%	0.48	1.75	100%	0.12	2.02
beta	32%	-0.04	0.11	53%	0.00	0.11	45%	0.00	0.09	28%	-0.02	0.13
absacc	10%	-0.01	0.14	9%	-0.03	0.12	9%	0.03	0.09	11%	-0.01	0.17
acc	11%	0.01	0.18	7%	0.02	0.20	6%	-0.06	0.13	9%	-0.02	0.19
aeavol	22%	0.01	0.12	20%	0.02	0.12	24%	0.00	0.07	20%	0.01	0.12
age	50%	-0.03	0.14	58%	-0.06	0.19	63%	-0.03	0.18	62%	-0.02	0.15
baspread	81%	-0.30	0.32	65%	-0.28	0.34	58%	-0.11	0.30	80%	-0.23	0.35
betasq	34%	-0.06	0.21	39%	0.01	0.24	55%	-0.04	0.17	35%	-0.03	0.19
bm_ia	38%	0.01	0.11	58%	0.01	0.13	46%	-0.01	0.10	43%	-0.02	0.12
cash	98%	0.02	0.17	92%	0.01	0.18	96%	0.07	0.13	99%	0.07	0.19
cashdebt	19%	0.20	0.20	18%	0.17	0.28	20%	0.05	0.23	20%	0.12	0.22
cashpr	100%	-0.02	0.17	100%	-0.02	0.20	100%	0.02	0.15	100%	0.05	0.20
cfp	56%	0.01	0.19	64%	-0.01	0.22	67%	-0.01	0.15	47%	-0.01	0.22
cfp_ia	21%	0.04	0.11	26%	0.03	0.12	28%	0.03	0.09	24%	0.04	0.11
chatoia	8%	-0.01	0.10	7%	-0.04	0.10	10%	-0.01	0.09	6%	0.00	0.10
chcsho	62%	0.01	0.10	50%	0.02	0.11	59%	-0.01	0.07	54%	-0.02	0.10
chempia	18%	0.02	0.11	17%	0.02	0.12	24%	0.00	0.08	15%	0.03	0.13
chinv	16%	0.01	0.10	10%	0.01	0.11	13%	0.02	0.08	11%	0.02	0.12
chpmia	38%	0.00	0.09	38%	0.01	0.11	45%	0.00	0.07	31%	0.01	0.10
chtx	29%	0.01	0.09	29%	0.03	0.11	34%	0.01	0.07	24%	0.00	0.10
cinvest	10%	0.00	0.08	6%	0.00	0.09	9%	0.00	0.07	6%	-0.02	0.09
convind	14%	0.00	0.07	11%	0.00	0.08	11%	0.00	0.04	17%	0.01	0.09
currat	31%	0.02	0.20	25%	0.08	0.22	21%	0.00	0.13	30%	0.06	0.22
depr	32%	0.03	0.13	41%	0.05	0.16	47%	0.08	0.10	24%	0.03	0.13
divi	8%	0.01	0.15	4%	-0.01	0.14	6%	-0.01	0.16	5%	-0.06	0.18
divo	0%	-0.05	0.31	0%	-0.02	0.15	1%	-0.29	0.47	1%	0.09	0.11
ear	16%	0.01	0.09	11%	0.02	0.09	18%	0.01	0.07	16%	-0.02	0.09
egr	41%	0.05	0.13	36%	0.06	0.15	54%	0.01	0.11	42%	0.03	0.13
ep	26%	0.08	0.22	58%	0.07	0.24	55%	0.00	0.17	25%	-0.02	0.23
gma	71%	0.02	0.19	74%	0.06	0.22	86%	0.10	0.16	63%	0.04	0.22
grcapx	17%	0.03	0.12	16%	0.03	0.12	29%	0.02	0.09	21%	0.03	0.11
grltnoa	22%	-0.01	0.12	15%	0.01	0.14	17%	-0.03	0.09	20%	0.01	0.15
herf	12%	0.06	0.13	25%	0.07	0.14	35%	0.06	0.11	19%	0.14	0.17
hire	32%	0.06	0.12	25%	0.06	0.14	26%	0.04	0.10	33%	0.03	0.13
idiovol	32%	-0.29	0.33	37%	-0.38	0.41	31%	-0.29	0.34	35%	-0.27	0.40
indmom	43%	0.03	0.10	42%	0.04	0.12	47%	0.03	0.09	42%	0.01	0.11
invest	9%	0.03	0.14	8%	0.03	0.14	12%	0.00	0.10	9%	0.05	0.12
lev	30%	-0.01	0.28	53%	0.00	0.31	58%	-0.07	0.27	31%	-0.03	0.33
lgr	46%	0.04	0.11	38%	0.02	0.12	44%	0.03	0.09	53%	0.05	0.11
mm12m	100%	0.15	0.21	100%	0.16	0.28	100%	0.10	0.22	100%	0.16	0.22
mom36m	98%	0.12	0.17	97%	0.10	0.20	100%	0.06	0.15	97%	0.12	0.19
ms	96%	0.06	0.12	91%	0.05	0.15	88%	0.03	0.10	96%	0.03	0.13
mve_ia	100%	-0.18	0.21	100%	-0.18	0.24	100%	-0.12	0.24	100%	-0.17	0.24
nincr	8%	0.00	0.07	6%	0.00	0.07	8%	0.00	0.05	5%	-0.01	0.09
orgcap	16%	-0.07	0.20	15%	-0.01	0.23	24%	0.06	0.19	18%	-0.05	0.22
pchcapx_ia	19%	-0.01	0.10	18%	-0.02	0.11	26%	-0.02	0.07	17%	-0.01	0.10
pchcurrat	2%	0.02	0.15	2%	0.02	0.10	2%	0.00	0.05	2%	-0.03	0.05
pchdepr	6%	-0.02	0.11	3%	0.00	0.08	3%	-0.01	0.06	4%	0.03	0.11
pchgm_pchsale	15%	0.02	0.10	12%	0.02	0.12	23%	0.01	0.07	13%	-0.02	0.13
pchquick	2%	0.01	0.15	2%	-0.01	0.09	3%	-0.01	0.06	2%	-0.01	0.09
pchsale_pchinvt	5%	0.02	0.32	3%	0.01	0.13	4%	0.09	0.21	2%	0.08	0.19
pchsale_pchrect	6%	-0.01	0.10	5%	-0.02	0.10	7%	-0.01	0.05	7%	0.02	0.10
pchsale_pchxsga	14%	0.00	0.11	9%	-0.01	0.13	11%	0.01	0.09	14%	0.01	0.12
pchsaleinv	4%	-0.01	0.34	3%	0.02	0.14	4%	-0.06	0.21	3%	0.05	0.14
pctacc	35%	-0.01	0.13	22%	-0.03	0.14	33%	0.04	0.11	33%	0.02	0.14
pricedelay	28%	0.00	0.11	31%	-0.01	0.12	33%	0.02	0.09	31%	0.00	0.12
ps	7%	0.01	0.12	3%	0.03	0.14	9%	0.00	0.05	7%	0.00	0.12
quick	32%	0.04	0.19	19%	0.03	0.22	27%	-0.01	0.13	26%	-0.03	0.23
rd	13%	0.00	0.07	7%	0.00	0.09	8%	0.04	0.07	12%	-0.02	0.07
rd_mve	18%	-0.12	0.19	17%	-0.08	0.20	17%	-0.03	0.14	18%	-0.09	0.17
rd_sale	16%	0.01	0.19	25%	0.01	0.19	23%	0.06	0.18	15%	0.12	0.23
realestate	41%	-0.02	0.15	31%	0.01	0.16	37%	-0.01	0.12	45%	0.02	0.15
roaq	42%	0.13	0.24	26%	0.20	0.28	27%	0.08	0.18	33%	0.11	0.24
roavol	8%	-0.05	0.15	12%	-0.04	0.16	10%	-0.05	0.12	6%	-0.03	0.17
roeq	75%	-0.01	0.19	90%	-0.01	0.20	84%	0.02	0.13	77%	0.00	0.19
roic	54%	0.13	0.18	52%	0.13	0.21	36%	0.08	0.17	48%	0.12	0.17
rsup	34%	0.05	0.13	37%	0.05	0.14	41%	0.03	0.09	31%	0.02	0.13
salecash	32%	-0.03	0.21	28%	-0.03	0.23	29%	-0.09	0.15	32%	-0.13	0.24
saleinv	13%	0.02	0.11	7%	-0.04	0.12	8%	-0.01	0.06	8%	0.00	0.14
salerec	6%	-0.02	0.14	8%	-0.05	0.14	16%	-0.04	0.12	9%	-0.01	0.15
secured	77%	0.05	0.19	83%	0.03	0.22	77%	0.01	0.11	72%	0.07	0.14
securedind	4%	0.01	0.18	1%	-0.04	0.30	3%	0.01	0.03	1%	0.00	0.04
sgr	43%	0.04	0.14	54%	0.04	0.15	58%	0.03	0.11	43%	0.06	0.14
sin	17%	0.02	0.15	13%	0.07	0.18	16%	0.05	0.14	12%	0.11	0.16
sp	95%	0.03	0.21	99%	0.05	0.27	96%	0.03	0.19	95%	-0.02	0.22
std_turn	65%	0.04	0.22	64%	0.03	0.23	76%	-0.05	0.16	70%	0.05	0.22
stdacc	37%	-0.03	0.16	16%	-0.06	0.17	33%	-0.05	0.17	34%	-0.04	0.19
stdcf	19%	-0.11	0.21	13%	-0.07	0.19	26%	-0.07	0.17	16%	-0.15	0.21
tang	11%	0.03	0.13	8%	0.00	0.14	15%	0.02	0.08	14%	0.06	0.11
tb	13%	0.04	0.11	13%	0.03	0.10	14%	0.00	0.06	11%	0.04	0.11
turn	67%	0.03	0.40	73%	0.13	0.41	78%	0.02	0.33	66%	0.08	0.40
zerotrade	56%	-0.12	0.47	77%	-0.08	0.45	59%	0.01	0.36	60%	-0.06	0.49
cons	100%	-2.20	1.08	100%	-2.86	1.55	100%	-4.86	1.83	100%	-2.57	1.67
Observations	169,177			37,746			5,503			11,046		

(Continued from the previous page)

Table 2.3: Variance Decomposition of Cross-Sectional Stock Returns

The cross-sectional variance of annual stock returns is decomposed into supply-side and demand-side effects, using equation (2.15). Supply-side effects include changes in shares outstanding, stock characteristics, and dividend yield. Demand-side effects include changes in AUM, coefficients on characteristics, and latent demand. The effects of changes in stock characteristics are further decomposed into a series of price changes induced by changing one stock characteristic at a time. Equal-weight and value-weight refer to the aggregation of elements in vector $\mathrm{Var}(\mathbf{r}_{t+1})$. Heteroskedasticity-robust standard errors are reported in Columns *S.E.*

	Equal-Weighted		Value-Weighted	
	% of Var	S.E.	% of Var	S.E.
Supply:				
Shares outstanding	2.1	0.2	1.5	0.5
Stock characteristics				
LNbe	4.3	0.1	1.3	0.5
profit	0.6	0.0	0.3	0.1
Gat	-0.2	0.0	0.1	0.1s
divA_be	0.2	0.1	-0.6	0.4
beta	0.4	0.0	0.5	0.1
absacc	0.0	0.0	0.0	0.0
acc	0.0	0.0	-0.1	0.1
aeavol	0.0	0.0	0.0	0.0
age	-0.1	0.0	-0.1	0.0
baspread	1.1	0.0	1.3	0.1
betasq	-0.1	0.0	-0.3	0.1
bm_ia	0.0	0.0	0.4	0.2
cash	0.2	0.0	0.1	0.1
cashdebt	0.0	0.0	-0.1	0.0
cashpr	0.1	0.0	0.2	0.1
cfp	0.0	0.0	0.2	0.1
cfp_ia	0.0	0.0	-0.1	0.1
chatoia	0.0	0.0	0.0	0.0
chcsho	-0.1	0.0	0.0	0.1
chempia	0.0	0.0	0.0	0.0
chinv	0.0	0.0	0.0	0.0
chpmia	0.0	0.0	-0.1	0.0
chtx	0.1	0.0	0.0	0.1
cinvest	0.0	0.0	0.0	0.0
convind	0.0	0.0	0.0	0.2
currat	-0.1	0.0	0.1	0.1
depr	0.0	0.0	0.0	0.0
divi	0.0	0.0	0.0	0.0
divo	0.0	0.0	0.0	0.0
ear	0.1	0.0	0.1	0.0
egr	0.0	0.0	0.1	0.0
ep	0.0	0.0	0.2	0.1
gma	0.0	0.0	0.0	0.1
grcapx	0.0	0.0	0.0	0.0
grltnoa	0.0	0.0	0.0	0.0
herf	0.0	0.0	-0.1	0.1
hire	0.0	0.0	0.0	0.1
idiovol	-0.3	0.0	-0.5	0.4
indmom	1.8	0.0	3.9	0.3
invest	0.0	0.0	0.0	0.0
lev	-0.1	0.0	-0.1	0.0
lgr	0.0	0.0	0.0	0.1

(Continued on the next page)

Table 2.3 (Continued): Variance Decomposition of Cross-Sectional Stock Returns

	Equal-Weighted		Value-Weighted	
	% of Var	S.E.	% of Var	S.E.
Supply :				
Stock characteristics				
mom12m	5.7	0.1	17.1	0.6
mom36m	0.0	0.0	0.1	0.2
ms	-0.1	0.0	0.2	0.1
mve_ia	0.0	0.0	-0.1	0.0
nincr	0.0	0.0	0.0	0.0
orgcap	0.0	0.0	0.0	0.0
pchcapx_ia	0.0	0.0	0.0	0.0
pchcurrat	0.0	0.0	0.0	0.0
pchdepr	0.0	0.0	0.0	0.0
pchgm_pchsale	0.0	0.0	0.0	0.0
pchquick	0.0	0.0	0.0	0.0
pchsale_pchinvt	0.0	0.0	0.0	0.0
pchsale_pchrect	0.0	0.0	-0.1	0.0
pchsale_pchxsga	0.0	0.0	-0.1	0.0
pchsaleinv	0.0	0.0	0.0	0.0
pctacc	0.1	0.0	0.0	0.1
pricedelay	0.0	0.0	0.1	0.0
ps	0.0	0.0	0.0	0.0
quick	-0.1	0.0	0.1	0.1
rd	0.0	0.0	0.0	0.0
rd_mve	0.0	0.0	0.1	0.0
rd_sale	0.0	0.0	-0.1	0.1
realestate	0.0	0.0	0.0	0.0
roaq	0.4	0.0	0.7	0.1
roavol	0.0	0.0	0.1	0.0
roeq	-0.2	0.0	0.0	0.1
roic	0.0	0.0	0.0	0.1
rsup	0.2	0.0	0.1	0.1
salecash	0.0	0.0	0.0	0.0
saleinv	0.0	0.0	0.0	0.0
salerec	0.0	0.0	0.0	0.0
secured	0.0	0.0	0.0	0.1
securedind	0.0	0.0	0.0	0.0
sgr	0.0	0.0	0.0	0.1
sin	0.0	0.0	0.0	0.0
sp	0.3	0.1	0.1	0.2
std_turn	-0.3	0.0	-0.3	0.2
stdacc	0.0	0.0	0.0	0.0
stdcf	0.0	0.0	0.0	0.0
tang	0.0	0.0	0.0	0.0
tb	0.0	0.0	0.0	0.0
turn	-0.9	0.2	0.4	0.6
zerotrade	1.2	0.1	0.5	0.2
Dividend yield	0.4	0.0	0.0	0.1
Demand:				
Assets under management	2.2	0.1	5.3	0.2
Coefficients on characteristics	5.8	0.3	14.1	1.0
Latent demand: extensive margin	22.6	0.3	21.6	1.3
Latent demand: intensive margin	52.8	0.4	31.8	1.8
Observations				134,278

(Continued from the previous page)

REFERENCES

Alberto Abadie. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2):391–425, 2021.

Zohair Alam, Adrian Alter, Jesse Eiseman, R. G. Gelos, Heedon Kang, Machiko Narita, Erlend Nier, and Naixi Wang. Digging Deeper–Evidence on the Effects of Macroprudential Policies from a New Database. Working Paper 2019/066, International Monetary Fund, 2019.

Felipe E. Aldunate, Zhi Da, Borja Larrain, and Clemens Sialm. Non-Fundamental Flows and Foreign Exchange Rates. Working Paper 30753, National Bureau of Economic Research, 2022.

Clifford S. Asness, Tobias J. Moskowitz, and Lasse Heje Pedersen. Value and Momentum Everywhere. *Journal of Finance*, 68(3):929–985, 2013.

Clifford S. Asness, Andrea Frazzini, and Lasse Heje Pedersen. Quality minus junk. *Review of Accounting Studies*, 24(1):34–112, 2019.

Jushan Bai. Panel Data Models with Interactive Fixed Effects. *Econometrica*, 77(4):1229–1279, 2009.

Jushan Bai and Serena Ng. Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 70(1):191–221, 2002.

Suman S. Basu, Emine Boz, Gita Gopinath, Francisco Roch, and Filiz D. Unsal. A Conceptual Model for the Integrated Policy Framework. Working Paper 2020/121, International Monetary Fund, 2020.

Alexandre Belloni, Daniel Chen, Victor Chernozhukov, and Christian Hansen. Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. *Econometrica*, 80(6):2369–2429, 2012.

Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. Inference on Treatment Effects after Selection among High-Dimensional Controls. *Review of Economic Studies*, 81(2):608–650, 2014.

Paula Beltran and Chang He. Inelastic Financial Markets and Foreign Exchange Interventions. Job market paper, UCLA, 2024.

Katharina Bergant, Prachi Mishra, and Raghuram Rajan. Cross-border Spillovers: How US Financial Conditions affect M&As Around the World. Working Paper 31235, National Bureau of Economic Research, 2022.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–890, 1995.

Balduin Bippus, Simon Lloyd, and Daniel Ostry. Granular Banking Flows and Exchange-Rate Dynamics. Working Paper 1043, Bank of England, 2024.

BIS. OTC foreign exchange turnover in April 2022. `https://www.bis.org/statistics/rpfx22_fx.pdf`, 2022. Accessed: 2013-10-16.

Nina Boyarchenko, Thomas M. Eisenbach, Pooja Gupta, Or Shachar, and Peter Van Tassel. Bank-Intermediated Arbitrage. Staff Report 858, Federal Reserve Bank of New York, 2020.

Fernando Broner, Alberto Martin, Lorenzo Pandolfi, and Tomas Williams. Winners and losers from sovereign debt inflows. *Journal of International Economics*, 130:103446, 2021.

Nelson Camanho, Harald Hau, and Hélene Rey. Global Portfolio Rebalancing and Exchange Rates. *Review of Financial Studies*, 35(11):5228–5274, 2022.

Huaizhi Chen. Cash Induced Demand. *Journal of Financial and Quantitative Analysis*, 59 (1):195–220, 2024.

Victor Chernozhukov, Christian Hansen, and Martin Spindler. Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments. *American Economic Review*, 105(5):486–490, 2015.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*, 21(1):C1–C68, 2018.

K. J. Martijn Cremers and Antti Petajisto. How Active Is Your Fund Manager? A New Measure That Predicts Performance. *Review of Financial Studies*, 22(9):3329–3365, 2009.

K. J. Martijn Cremers, Miguel A. Ferreira, Pedro Matos, and Laura Starks. Indexing and active fund management: International evidence. *Journal of Financial Economics*, 120 (3):539–560, 2016.

J. Bradford De Long, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Noise Trader Risk In Financial Markets. *Journal of Political Economy*, 98(4):703–738, 1990.

Lisa De Simone, Joseph D. Piotroski, and Rimmy E. Tomy. Repatriation Taxes and Foreign Cash Holdings: The Impact of Anticipated Tax Reform. *Review of Financial Studies*, 32 (8):3105–3143, 2019.

Wenxin Du and Jesse Schreger. CIP deviations, the dollar, and frictions in international capital markets. In *Handbook of International Economics*, volume 6, pages 147–197. Elsevier, 2022.

Wenxin Du, Alexander Tepper, and Adrien Verdelhan. Deviations from Covered Interest Rate Parity. *Journal of Finance*, 73(3):915–957, 2018.

Wenxin Du, Benjamin Hébert, and Amy Wang Huber. Are Intermediary Constraints Priced? *Review of Financial Studies*, 36(4):1464–1507, 2023.

Darrell Duffie. Post-Crisis Bank Regulations and Financial Market Liquidity. Paolo Baffi Lecture on Money and Finance, Bank of Italy, 2017.

Martin D. D. Evans and Richard K. Lyons. Order Flow and Exchange Rate Dynamics. *Journal of Political Economy*, 110(1):170–180, 2002.

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

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

Eugene F. Fama and Kenneth R. French. Choosing factors. *Journal of Financial Economics*, 128(2):234–252, 2018.

Xiang Fang and Yang Liu. Volatility, intermediaries, and exchange rates. *Journal of Financial Economics*, 141(1):217–233, 2021.

Yusheng Fei. *The Year-End Effect in the Foreign Exchange Market*. PhD thesis, University of Chicago, 2023.

Guanhao Feng, Stefano Giglio, and Dacheng Xiu. Taming the Factor Zoo: A Test of New Factors. *Journal of Finance*, 75(3):1327–1370, 2020.

Joachim Freyberger, Andreas Neuhierl, and Michael Weber. Dissecting Characteristics Nonparametrically. *Review of Financial Studies*, 33(5):2326–2377, 2020.

FTSE. Global Equity Index Series Guide to Calculation v3.5. `https://www.lseg.com/content/dam/ftse-russell/en_us/documents/ground-rules/ftse-global-equity-index-series-guide-to-calc.pdf`, 2023. Accessed: 2013-10-15.

Xavier Gabaix. Behavioral inattention. In *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 2, pages 261–343. Elsevier, 2019.

Xavier Gabaix and Ralph S. J. Koijen. In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis. Working Paper 28967, National Bureau of Economic Research, 2021.

Xavier Gabaix and Ralph S. J. Koijen. Granular Instrumental Variables. Working Paper 28204, National Bureau of Economic Research, 2022.

Xavier Gabaix and Matteo Maggiori. International Liquidity and Exchange Rate Dynamics. *Quarterly Journal of Economics*, 130(3):1369–1420, 2015.

Xavier Gabaix, Ralph S. J. Koijen, Robert Richmond, and Motohiro Yogo. Asset Embeddings. Working paper, University of Chicago, 2024.

Nicolae Garleanu and Lasse Heje Pedersen. Margin-based Asset Pricing and Deviations from the Law of One Price. *Review of Financial Studies*, 24(6):1980–2022, 2011.

Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8):2586–2624, 2020.

Pierre-Olivier Gourinchas, Walker D. Ray, and Dimitri Vayanos. A Preferred-Habitat Model of Term Premia, Exchange Rates, and Monetary Policy Spillovers. Working Paper 29875, National Bureau of Economic Research, 2022.

Jeremiah Green, John R. M. Hand, and X. Frank Zhang. The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. *Review of Financial Studies*, 30(12):4389–4436, 2017.

Robin Greenwood, Samuel G. Hanson, Jeremy C. Stein, and Adi Sunderam. A Quantity-Driven Theory of Term Premia and Exchange Rates. *Quarterly Journal of Economics*, 138(4):2327–2389, 2023.

Denis Gromb and Dimitri Vayanos. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics*, 66(2-3):361–407, 2002.

Shihao Gu, Bryan T. Kelly, and Dacheng Xiu. Empirical Asset Pricing via Machine Learning. *Review of Financial Studies*, 33(5):2223–2273, 2020.

Michelle Hanlon, Rebecca Lester, and Rodrigo Verdi. The effect of repatriation tax costs on U.S. multinational investment. *Journal of Financial Economics*, 116(1):179–196, 2015.

Samuel M. Hartzmark and David H. Solomon. Predictable Price Pressure. Working Paper 30688, National Bureau of Economic Research, 2022.

Harald Hau. Global versus Local Asset Pricing: A New Test of Market Integration. *Review of Financial Studies*, 24(12):3891–3940, 2011.

Harald Hau and Helene Rey. Exchange Rates, Equity Prices, and Capital Flows. *Review of Financial Studies*, 19(1):273–317, 2006.

Harald Hau, Massimo Massa, and Joel Peress. Do Demand Curves for Currencies Slope Down? Evidence from the MSCI Global Index Change. *Review of Financial Studies*, 23 (4):1681–1717, 2010.

Zhiguo He and Arvind Krishnamurthy. Intermediary Asset Pricing. *American Economic Review*, 103(2):732–770, 2013.

Zhiguo He, Bryan T. Kelly, and Asaf Manela. Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35, 2017.

Ethan Ilzetzki, Carmen M. Reinhart, and Kenneth S. Rogoff. Exchange Arrangements Entering the Twenty-First Century: Which Anchor will Hold? *Quarterly Journal of Economics*, 134(2):599–646, 2019.

Oleg Itskhoki and Dmitry Mukhin. Exchange Rate Disconnect in General Equilibrium. *Journal of Political Economy*, 129(8):2183–2232, 2021.

Zhengyang Jiang, Robert J. Richmond, and Tony Zhang. A Portfolio Approach to Global Imbalances. Working Paper 30253, National Bureau of Economic Research, 2022.

Wei Jiao, G. Andrew Karolyi, and David T. Ng. Can International Funds Navigate Changing Global Investment Environments? Working paper, Cornell University, 2024.

Hyeyoon Jung. Real Consequences of Shocks to Intermediaries Supplying Corporate Hedging Instruments. Staff Report 989, Federal Reserve Bank of New York, 2023.

Bryan T. Kelly, Seth Pruitt, and Yinan Su. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524, 2019.

Ralph S. J. Koijen and Motohiro Yogo. A Demand System Approach to Asset Pricing. *Journal of Political Economy*, 127(4):1475–1515, 2019.

Ralph S. J. Koijen and Motohiro Yogo. Exchange Rates and Asset Prices in a Global Demand System. Working Paper 27342, National Bureau of Economic Research, 2024.

Ralph S. J. Koijen, Robert Richmond, and Motohiro Yogo. Which Investors Matter for Equity Valuations and Expected Returns? Working Paper 27402, National Bureau of Economic Research, 2023.

Naz Koont, Yiming Ma, Lubos Pastor, and Yao Zeng. Steering a Ship in Illiquid Waters: Active Management of Passive Funds. Working Paper 30039, National Bureau of Economic Research, 2023.

Pentti J. K. Kouri. The Exchange Rate and the Balance of Payments in the Short Run and in the Long Run: A Monetary Approach. *Scandinavian Journal of Economics*, 78(2): 280–304, 1976.

Serhiy Kozak, Stefan Nagel, and Shrihari Santosh. Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292, 2020.

Jiacui Li. What Drives the Size and Value Factors? *Review of Asset Pricing Studies*, 12(4): 845–885, 2022.

Andrew Lilley, Matteo Maggiori, Brent Neiman, and Jesse Schreger. Exchange Rate Reconnect. *Review of Economics and Statistics*, pages 1–28, 2019.

Andrew W. Lo. The Statistics of Sharpe Ratios. *Financial Analysts Journal*, 58(4):36–52, 2002.

Dong Lou. A Flow-Based Explanation for Return Predictability. *Review of Financial Studies*, 25(12):3457–3489, 2012.

Hanno Lustig, Nikolai Roussanov, and Adrien Verdelhan. Common Risk Factors in Currency Markets. *Review of Financial Studies*, 24(11):3731–3777, 2011.

Richard K. Lyons. *The Microstructure Approach to Exchange Rates*. The MIT Press, 2001.

Matteo Maggiori. International macroeconomics with imperfect financial markets. In *Handbook of International Economics*, volume 6, pages 199–236. Elsevier, 2022.

Matteo Maggiori, Brent Neiman, and Jesse Schreger. International Currencies and Capital Allocation. *Journal of Political Economy*, 128(6):2019–2066, 2020.

Lukas Menkhoff, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. Currency momentum strategies. *Journal of Financial Economics*, 106(3):660–684, 2012.

MSCI. Index Calculation Methodology, August 2023. `https://www.msci.com/index/methodology/latest/IndexCalc/`, 2023. Accessed: 2013-10-15.

Lorenzo Pandolfi and Tomas Williams. Capital flows and sovereign debt markets: Evidence from index rebalancings. *Journal of Financial Economics*, 132(2):384–403, 2019.

Claudio Raddatz, Sergio L. Schmukler, and Tomás Williams. International asset allocations and capital flows: The benchmark effect. *Journal of International Economics*, 108:413–430, 2017.

Stefan Reitz and Dennis Umlandt. Currency returns and FX dealer balance sheets. *Journal of International Economics*, 133:103541, 2021.

Mirela Sandulescu, Fabio Trojani, and Andrea Vedolin. Model-Free International Stochastic Discount Factors. *Journal of Finance*, 76(2):935–976, 2021.

Simon Niklas Maria Schmickler. *Essays on Institutional Trading and Liquidity*. PhD thesis, Princeton University, 2022.

Kurt Schmidheiny and Sebastian Siegloch. On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications. Discussion Paper 13477, Center for Economic and Policy Research, 2020.

Andrei Shleifer and Robert W. Vishny. The Limits of Arbitrage. *Journal of Finance*, 52(1): 35–55, 1997.

Clemens Sialm and Qifei Zhu. Currency Management by International Fixed Income Mutual Funds. Working Paper 29082, National Bureau of Economic Research, 2021.

Yiuman Tse. Return seasonality in the foreign exchange market. *Applied Economics Letters*, 25(1):5–8, 2018.

Adrien Verdelhan. The Share of Systematic Variation in Bilateral Exchange Rates. *Journal of Finance*, 73(1):375–418, 2018.

Martin J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Number 48 in Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

Scott M. Weiner. *The Complete Guide to ETF Portfolio Management: The Essential Toolkit for Practitioners*. McGraw Hill, 2023.

Peng Zhao and Bin Yu. On Model Selection Consistency of Lasso. *Journal of Machine Learning Research*, 7:2541–2563, 2006.