EPFL

Thèse n° 10 116

Demand-based Asset Pricing: Theory, Estimation and Applications

Présentée le 29 juin 2023

Collège du management de la technologie
Chaire du Prof. ordinaire Collin-Dufresne
Programme doctoral en finance

pour l'obtention du grade de Docteur ès Sciences

par

Philippe VAN DER BECK

Acceptée sur proposition du jury

Prof. E. Morellec, président du jury
Prof. P. Collin Dufresne, directeur de thèse
Prof. M. Yogo, rapporteur
Prof. L. Bretscher, rapporteur
Prof. S. Malamud, rapporteur

École
polytechnique
fédérale
de Lausanne

2023

Acknowledgements

I would like to begin by expressing my deepest gratitude to my advisor, Pierre Collin-Dufresne, who has been an incredible source of guidance, encouragement, and support throughout my PhD. Your invaluable advice and expertise have been instrumental in shaping my research and I am truly fortunate to have had you as my mentor. I would also like to thank Motohiro Yogo for making my visiting semester at Princeton University possible, for generously devoting his time and attention to me during this amazing semester, and for his invaluable feedback on my research. I am also indebted to Lorenzo Bretscher for his incredible support during the job market.

Most importantly, I would like to express my heartfelt thanks to my mom, dad, and sister for their unwavering love and support. I would also like to thank my aunt Sibille. I am forever grateful for your contribution to my academic and personal growth. Finally, I would like to sincerely thank Sebastian & Achmed Hasslinger, Max Sauer, Maxwell Dean, Romain Gratis, Ole Gwakam, Pierre Cailloux, Jonathan Abruzzi, Eric Gunshof, Bula, Yercomo Melegati, Birgit Siefert, Fred Sandig, Mölle, Kendrick Depay, Olivier Krek, Sam Wagner, and all the members of the Club der Dichteren for their insightful feedback on my research.

Lausanne, April 21, 2023 P. vdB.

Abstract

This thesis investigates the relationship between investors' demand shocks and asset prices through the use of data on portfolio holdings. In three chapters, I study the theory, estimation, and application of demand-based asset pricing models, which incorporate data on investors' portfolio holdings and equilibrium asset prices. I first present a generalized framework and propose a new estimator of investor-specific demand curves that is based on time-series changes in investors' portfolios. I then use and extend the proposed estimator to quantify the equilibrium price impact of the growing institutional demand for sustainable investments. I show that the returns from sustainable investing are strongly driven by price pressure from flows towards sustainable funds, causing high realized returns that do not reflect high expected returns. The last chapter quantifies the price impact of the retail investment boom during the Covid-19 pandemic via a structural model that uses data on portfolio holdings of US households.

Zusammenfassung

Diese Arbeit untersucht den Zusammenhang zwischen Vermögenspreisen und der Nachfragefunktion von Investoren mittels Daten zu Portfolio-positionen. In drei Kapiteln werden Theorie, Schätzung und Anwendung von nachfragebasierten Vermögenspreismodellen untersucht. Diese Modelle schätzen die Nachfragefunktion institutioneller Investoren und verknüpfen diese mit Vermögenspreisen. Zunächst wird ein generalisiertes Modell vorgestellt und ein neuer Schätzer für investorenspezifische Nachfragekurven vorgeschlagen, der auf zeitlichen Veränderungen in den Portfolio-positionen der Investoren basiert. Anschließend wird der vorgeschlagene Schätzer verwendet und erweitert, um den Einfluss der wachsenden institutionellen Nachfrage nach nachhaltigen Investitionen zu quantifizieren. Es wird gezeigt, dass die Renditen von nachhaltigen Investitionen stark von Nachfrageschocks getrieben sind, was zu hohen realisierten Renditen führt, die keine hohen erwarteten Renditen widerspiegeln. Das letzte Kapitel quantifiziert den Einfluss von Robinhood-Tradern auf die Aktienkurse während der Covid-19-Pandemie anhand eines strukturellen Modells, das Daten zu den Portfolio-positionen von US Haushalten verwendet.

Contents

Introduction

This thesis investigates the relationship between investors' demand shocks and equilibrium asset prices through the use of data on the asset holdings of different investors. If the aggregate demand for financial securities is downward sloping, then data on investors' portfolio choice potentially contains valuable information on what drives equilibrium asset prices. Following the influential work of Koijen and Yogo (2019), a rapidly growing literature links holdings data and equilibrium asset prices via structural models of investor demand. These models have since been commonly referred to as "demand-based asset pricing models". In three chapters, I study the theory, estimation and application of such demand-based asset pricing models.

The first chapter sets the stage with a generalized theoretical framework that links investor demand and asset prices and nests many workhorse asset pricing models. The key parameter in relating demand and equilibrium prices is investors' elasticity of demand with respect to the price. Unlike previous studies, which rely on cross-sectional estimates in levels, this chapter proposes estimating elasticities from investors' trades, that is changes in their portfolios. I use demand shocks from mutual fund flows as an instrument to address the endogeneity of trades and prices. Using the estimation in changes along with the flow-based instrument I find that elasticities are 4 times larger than what previous estimates suggest. Estimation over different trading horizons furthermore shows that investors become more elastic in the long run. The results suggest that the impact of demand shocks on equilibrium prices is smaller than previously estimated and partly reverts over time.

The second chapter employs the estimation technique proposed in the first chapter to quantify the equilibrium price impact of the growing institutional demand for sustainable investments. I show that the returns from sustainable investing are strongly driven by price pressure from flows towards sustainable funds, causing high realized returns that do not reflect high expected returns. Using a structural model, I estimate investors' ability to accommodate the demand from sustainable funds, which is given by their elasticity of substitution between stocks. I show that every dollar flowing from the market portfolio into sustainable mutual funds increases the aggregate value of green stocks by \$0.4. The price pressure from flows supports the effectiveness of impact investing by lowering green firms' cost of capital. In the absence of flow-driven price pressure, sustainable funds would have underperformed the market from 2016 to 2021. To this end, I develop a new measure of total capital flows into managed portfolios. The price pressure from total ESG flows is highly correlated with empirically observed returns, both in the time-series and in the cross-section. I support the structural estimates with reduced-form evidence, showing that index inclusions and mandate-driven portfolio additions by sustainable mutual funds significantly boost the prices of green stocks.

The last chapter is dedicated to the retail investment boom during the Covid-19 pandemic. In a joint paper with Coralie Jaunin, we quantify the impact of Robinhood traders on the US equity market. We estimate retail and institutional demand curves using a structural model and derive aggregate pricing implications via market clearing. The inelastic nature of institutional demand allows Robinhood traders to have a substantial effect on stock prices. Despite their negligible market share of 0.2%, Robinhood traders account for 10% of the cross-sectional variation in stock returns during the second quarter of 2020. Furthermore, without the surge in retail trading activity the aggregate market capitalization of the smallest size quintile of stocks would have been 25% lower.

Chapter 1

On the Estimation of Demand-based Asset Pricing Models

1.1 Introduction

A wealth of detailed data on investor's portfolio holdings has breathed new life into the question of whether changes in investor demand drive changes in prices. If the aggregate demand for a financial security is downward sloping, investors' non-fundamental demand shocks can have a meaningful impact on equilibrium prices. The Demand System Approach to Asset Pricing developed by Koijen and Yogo (henceforth KY, 2019) allows incorporating holdings data into equilibrium asset pricing by estimating investor-specific demand curves. Investors' elasticity of demand with respect to the price (henceforth elasticity of demand) lies at the heart of the demand system approach. It determines the equilibrium price impact of counterfactual experiments within the demand system. If investors' elasticity of demand is high, then equilibrium prices do not have to move a lot in order to accommodate demand shocks. Importantly, the notion of the elasticity of demand refers to how strongly investors *trade* in response to *changes* in prices.

Under valid identification, the estimated elasticities should not depend on the estimation specification. In contrast, this paper highlights, that estimates from changes in of portfolio holdings (i.e., how investors trade) are substantially different to estimates from levels (i.e., what investors hold). KY (2019), and all subsequent papers using their approach, infer price elasticities of demand from the cross-section of holdings thereby muting the time series dimension. In particular, they run cross-sectional regressions of portfolio weights onto prices using an instrument to control for their joint endogeneity. The underlying intuition is that investors holding a larger weight in cheaper securities (controlling for fundamental value) are more price-elastic than passive investors simply holding the market portfolio. Importantly, the estimation does not take into account how investors *trade*, i.e., how their holdings change over time. This constitutes an important deviation from the event-study evidence that identifies the slope of the aggregate demand curve by assessing the price *change* caused by *changes* in non-fundamental demand due to e.g. index inclusions or mutual fire sales. In this spirit, I propose estimating price elasticities of demand from quarterly changes in holdings (i.e., trades) as a response to non-fundamental price changes. The estimation using trades has several advantages. First, it removes the omitted variable bias from unobservable static portfolio tilts (such as investment mandates) that are correlated with the price. Second, it allows directly estimating investors' change in demand as a response to changing prices, as opposed to inferring it from the cross-section of holdings in levels. The identification in KY (2019) requires investor-level holdings across multiple securities. The identification proposed in this paper is not restricted to a multiple asset setting. In fact, estimating elasticities using investors' trades allows estimating demand systems for individual securities such as the aggregate stock market, single stocks, or factor portfolios. Third, the estimated elasticities are non-negative for all investors, which makes intuitive sense as (controlling for fundamental value) a higher price should lower demand. In contrast, the unconstrained cross-sectional identification from holdings leads to many negative elasticities, particularly for large investors. Thus KY (2019) have to impose a coefficient constraint in order to ensure unique equilibrium prices in their counterfactuals. Fourth, changes in holdings (i.e. trades) can be computed over an arbitrary time horizon. This allows estimating elasticities at different frequencies. The estimation in KY (2019) essentially ignores the frequency at which holdings data is available. In its current form, the demand system therefore implies that demand shocks have a permanent impact on prices because the price elasticities of demand do not have a time-series component. The time series nature of trades as opposed to holdings allows linking demand-based asset pricing to the return dynamics (e.g. long run reversal) of non-fundamental demand shocks. Fifth, by including time-series information, the estimation in changes allows identifying stock-specific elasticities as well as cross-elasticities. Sixth, one can directly apply existing instruments from the event-study literature such as mutual fund flows or index reconstitutions. The event studies can furthermore be used to verify counterfactual equilibrium prices obtained from the demand system approach. This lends additional flexibility to the demand-based approach. It is important to note, that any asset pricing model is *demand-based* and implies elasticity of demand with respect to the price that can in general be estimated from observable data on portfolio choice. Neoclassical models typically imply high price elasticities of demand such that demand shocks have minor equilibrium price effects (see the calibrations in Petajisto (2009) and Gabaix and Koijen (2021)). The negligible role of demand shocks in standard models is at odds with extensive empirical evidence from index inclusions (Wurgler and Zhuravskaya (2002)), mutual fund flows (Lou (2012)), index reconstitutions (Chang et al. (2015)), benchmarking (Pavlova and Sikorskaya (2022)) and dividend reinvestments (Hartzmark and Solomon (2021)) suggesting that investors have downward-sloping (or inelastic) demand curves. The notion of inelastic demand for financial securities is not new and has been around since at least Shleifer (1986). However, the increased availability of investor-level holdings data has revived the interest in estimating the slope of investors' demand curves and linking the holdings to equilibrium prices.

I start by framing the demand-based approach in a generalized form that nests many workhorse asset pricing frameworks. I show that the equilibrium price effects of counterfactual experiments within the demand system can be approximated as the product of a dollar demand shock and a multiplier matrix that is the inverse of the aggregate elasticity of demand. Correctly estimating elasticities (in particular their magnitude) therefore critically determines whether the demand-based approach produces realistic counterfactuals. Price elasticities of demand are defined as the percentage change in holdings as a response to a one percent increase in the price. The most natural estimation specification that directly emerges from the definition is to estimate investors' percentage changes in holdings as a response to non-fundamental price changes. This essentially corresponds to first-difference estimator of the demand curve in KY (2019). Using the same data, estimating their specification in first differences produces higher elasticities that are uncorrelated to the estimates in KY (2019). This points towards an omitted variable bias (due to e.g. unobservable investment mandates that are correlated with prices) in the cross-sectional identification using holdings. Identification via trades mitigates endogeneity concerns due to unobservable portfolio tilts. I confirm this bias in a simple simulation by estimating the demand of a mandated investor that is perfectly inelastic. Regardless of using holdings or changes in holdings (i.e., trades) in the estimation, a causal identification requires instruments to control for the joint endogeneity of prices and demand. Identifying the elasticity of demand to the price using holdings as in KY (2019) requires exogenous variation in prices, which is difficult to find. However, the literature provides a variety of different approaches to construct exogenous variation in returns (i.e., price changes), which can be used to identify elasticities via trades. Hypothetical trades based on mutual fund flows (see e.g. Edmans et al. (2012)) are ideally suited as an instrument as they are available for all stocks at all quarters. However, mutual fund flows contain fundamental information if they are driven by the funds' underlying portfolio tilts. I therefore construct flow-induced demand shocks unrelated to fundamental news by orthogonalizing flows with respect to fund-specific characteristics. The instrument is highly relevant, i.e., it is strongly related to contemporaneous returns controlling for known return predictors. The price pressure furthermore reverts after approximately 1.5 years. This is suggestive evidence that the demand shocks are non-fundamental (i.e., that the instrument is exogenous). I then proceed by estimating investor-specific elasticities using flow-induced mutual fund demand as an instrument. The resulting elasticities of demand to the price are 4 times larger than the original estimates by KY (2019). This has important consequences for the equilibrium prices obtained from counterfactual experiments, because the price impact of demand shocks is exclusively driven by the magnitude of the estimated elasticities. I then outline how the estimation using trades as opposed to holdings can i) be incorporated in existing demand-based frameworks, ii) produce stock-specific and cross-elasticites, iii) reveal high frequency estimates using alternative holdings data, iv) be used to obtain *causal* demand responses to other variables such as ESG scores.

Related Literature. This paper relates to two strands of literature. Most importantly, it directly relates to the recent and growing literature on demand-based asset pricing initiated by KY (2019). This literature uses data on investors' portfolio holdings for different asset classes to estimate demand curves and, via the price elasticities of demand, link them to equilibrium prices (e.g. Koijen et al. (2022), Koijen and Yogo (2020), Haddad et al. (2021), Han et al. (2021), Noh and Oh (2020), Benetton and Compiani (2020), Bretscher et al. (2020), Jiang et al. (2020), van der Beck and Jaunin (2021), Jansen (2021)). All of the above-mentioned papers use the demand system approach to conduct counterfactual experiments, the impact of which critically hinges on the elasticity estimates. The impact of the counterfactual experiments would be much smaller under the higher elasticity estimates identified via trades instead of holdings. Identifying elasticities via trades requires exogenous variation in price changes (i.e returns) as opposed to exogenous variation in price levels as in KY (2019). Thus this paper relates to the large literature on identifying non-fundamental price pressure from index reconstitutions (see e.g. Shleifer (1986), Wurgler and Zhuravskaya (2002), Chang et al. (2015), Pavlova and Sikorskaya (2022), Greenwood (2005)), mutual fund flows (see e.g. Edmans et al. (2012), Coval and Stafford (2007), Lou (2012), Wardlaw (2020), Schmickler (2020)) or dividend reinvestments (see e.g. Hartzmark and Solomon (2021)). All of the measures introduced in these papers can essentially be used as instruments to estimate elasticities from trades. This lends additional flexibility to the demand-based approach as the estimates can be verified in a host of different settings.

The remainder of this paper is structured as follows. Section 3.2 describes the role of price elasticities in asset pricing in a generalized setting. Section 3.3 shows how to identify price elasticities from holdings data. Section 3.4 is dedicated to identification. I describe how to construct a flow-based instrument that can be used to causally identify elasticities and compare my estimates to the estimates in KY (2019). In Section 3.5, I outline potential applications and extensions of the demand estimation using trades as opposed to holdings. Section 3.6 concludes.

1.2 The Role of Price Elasticities in Asset Pricing

The demand-based approach to asset pricing allows to jointly match data on portfolio holdings and equilibrium asset prices by estimating investor-specific demand curves. The key obstacle in estimating investors' demand curves is the identification of the elasticity of demand (i.e. the how elastically demand responds to a change in the price). Below I provide a generalized version of the demand system in KY (2019) and show that it nests many workhorse frameworks in asset pricing.

A Generalized Demand-Based Asset Pricing

There are N stocks indexed by $n = 1, ..., N$ and 2 time periods $t = 0, 1$. Shares outstanding of each stock are given by Q^*, normalized to 1 so that the endogenous total market equity of each stock is given by its price $P(n)$. Each stock pays an cashflow $D(n)$ in period 1 with a covariance matrix $\Sigma \in \mathbb{R}^{N\times N}$. There are I investors indexed by $i = 1, \ldots, I$ whose assets under management A^i are exogenously determined e.g. by the money that households allocate to them. This can be easily endogenized. The funds may invest in the N stock, as well as a riskless asset whose exogenous interest rate is normalized to 0. Let $Q^i(n)$ denote the number of shares held by i in stock n and $w^i(n) = \frac{Q^i(n)P(n)}{A^i}$ the corresponding portfolio weight. For now, I am assuming that the optimal portfolio $Q^i \in \mathbb{R}^N$ is some unknown fund-specific function $f^i(\cdot)$ of the vector of stock prices $P \in \mathbb{R}^N$ and a collection of other exogenous observable and unobservable variables X (such as the assets under management, interest rate, fundamentals or investment constraints):

$$Q^i = f^i\left(P, X\right) \tag{1.1}$$

For example, under CARA preferences with risk aversion equal to γ^i, the institution maximizes $\mathbb{E}[e^{-\gamma^i A_1^i}]$ subject to the budget constraint $A_1^i = A_0^i + Q^{i\prime}(\mathbb{E}[D(n)] - P)$ by choosing the optimal portfolio

$$Q^i = (\gamma^i \Sigma)^{-1}(\mathbb{E}[D(n)] - P). \tag{1.2}$$

The general specification (1.1) allows plugging in a demand curve which matches observable data on portfolio choice. For example KY (2019) use an exponential function for $f^i(\cdot)$ whereas Balasubramaniam et al. (2021) and Betermier et al. (2022) opt for a linear specification. Micro-founding a demand curve that matches observable data is beyond the scope of this paper and left for future research.[1] Importantly, $f^i(P, X)$ pins down an investor's elasticity of demand, which is given by

$$\zeta_n^i = -\frac{\partial Q^i(n)}{\partial P(n)} \frac{P(n)}{Q^i(n)}. \tag{1.3}$$

It measures how much of her holdings in stock n the investor sells (in %) when the price of n goes up by 1 %. Similarly, cross-elasticities are defined as $\zeta_{n,m}^i = -\frac{\partial Q^i(n)}{\partial P(m)} \frac{P(m)}{Q^i(n)}$ measuring how much of stock n the investor sells when the price of an different stock m goes up by 1 %. For example, assume that Apple's share price goes up by 1% due to some non-fundamental shock (such as an index reconstitution), causing Blackrock to sell 2% of its shares in Apple and substituting towards Google by increasing its shares in Google by 0.5%. This implies that the elasticity of Blackrock's demand for Apple with respect to the price of Apple is $\zeta_{\text{Apple}}^{\text{Blackrock}} = 2$, and that the elasticity of Blackrock's demand for Apple with respect to the price of Google is $\zeta_{\text{Google,Apple}}^{\text{Blackrock}} = -0.5$. Let ζ^i denote the $N \times N$ elasticity matrix that has price elasticities on the diagonal and the cross-price elasticities on the off-diagonal elements. In the CARA case above, the investor-specific elasticity matrix ζ^i is proportional to $(\gamma^i\Sigma)^{-1}$.[2] More risk averse investors require a larger price concession to move away from the mean-variance efficient portfolio and are hence more inelastic with respect to the price. Similarly, all investors are more inelastic with respect to riskier stocks as these stocks contribute more to the risk of

[1] KY (2019) provide a suggestive micro-foundation of log-linear portfolio weights by using the fact that $w(n) = 1 + X(n) + \frac{1}{2!}X(n)^2 \ldots \approx e^{X(n)}$.

[2] Formally, the elasticity in the CARA case is given by $\text{diag}(Q_t^i)^{-1}(\gamma^i\Sigma)^{-1}\text{diag}(P_t)$. Note, that in the CARA case it is more convenient to define elasticities in absolute terms (as opposed to percentages), i.e. $\zeta^i = \frac{\partial Q^i(n)}{\partial P(n)}$. In this case the elasticity matrix is given by $(\gamma^i\Sigma)^{-1}$

the arbitrage portfolio (see e.g. Greenwood (2005)). Note, that the frictionless CARA setting produces price elasticities of demand that are too large to be reconciled with the observed price impact of index inclusions (see the calibration in Petajisto (2009)). However, it serves as a useful benchmark on how to think about price elasticities in a classical framework.

Market clearing implies that all funds must jointly hold the market portfolio Q^*. Summing (1.1) across investors yields:

$$\sum_{i=1}^{I} f^i(P,X) = Q^* \tag{1.4}$$

In the CARA example above, equilibrium prices are available in closed form and given by the CAPM

$$P = \mathbb{E}[D] - \gamma \Sigma Q^* \tag{1.5}$$

where $\gamma \equiv 1/\left(\sum_{i=1}^{I} 1/\gamma^i\right)$ is the market's effective risk aversion. The exponential demand specification in KY (2019) does not allow closed-form equilibrium prices. The price effects of counterfactual experiments have to be computed numerically. The next section shows, however, shows that they can be approximated as the product of a dollar demand shock and the market's aggregate elasticity of demand.

B Demand Shocks and Equilibrium Prices

Now, assume that at time $t = 0$, there is shock ΔX that changes the demand of an investor resulting in a dollar trade of ΔD. The shock could be due to flows altering an investor's assets under management or portfolio rebalancing due to a change in stock-specific characteristics.

PROPOSITION 1 *A first order approximation of the equilibrium change in prices ΔP due to the demand shock ΔD is given by*

$$\Delta P = \mathcal{M} \Delta D \tag{1.6}$$

where $\mathcal{M} \in \mathbb{R}^{N\times N}$ is given by

$$\mathcal{M} = \left(\sum_{i=1}^{I} diag(Q)^i \zeta^i \right)^{-1} \tag{1.7}$$

See Appendix A for a proof. Proposition 1 states that the impact of demand shocks on the cross-section of realized returns is given by $\mathcal{M}$, which is the inverse of the aggregate elasticity of demand (weighted by ownership Q^i). $\mathcal{M}$ is the cross-sectional pendant to the scalar multiplier in Gabaix and Koijen (2021) and will henceforth be referred to as the multiplier matrix. It measures the market's willingness to substitute between stocks. The more price-elastic investors are (i.e., the larger the diagonal elements in ζ^i), the less prices have to move to accommodate the exogenous demand shock. Cross-price elasticities drive the off-diagonal elements in $\mathcal{M}$ and are responsible for spill-over effects to other stocks. For example, if investors accommodate an exogenous demand shock for green stocks primarily by substituting towards brown industries, the relative price impact of ESG investing may be negligible. Ultimately, elasticities ζ^i and therefore the multiplier matrix $\mathcal{M}$ are pinned down by the specific choice of $f^i(P, X)$. In the CARA case the most intuitive way to model exogenous demand ΔD is via a change in supply Q^*, such as the inclusion of some stocks in an index. The resulting price change is given by $\Delta P = \gamma\Sigma\Delta Q^*$. Thus the CARA setting implies that the multiplier matrix is proportional to $\gamma\Sigma$. It is driven by the covariance of cashflows and the market's effective risk aversion. Riskier stocks have a larger multiplier because they contribute more to the arbitrage portfolio (see e.g. Greenwood (2005)). Note, that the channel through which demand shocks affect equilibrium prices in the CAPM is entirely risk-driven: If investors are risk neutral, or future cashflows are risk-free then $P = \mathbb{E}[D]$ and exogenous demand shocks do not affect prices.

C Reevaluating Counterfactual Experiments

KY (2019) show how to compute $\mathcal{M}$ by estimating investor-specific elasticities ζ^i from holdings data. In their model, portfolio weights are exponential linear in prices and stock-specific characteristics.[3] Because of the non-linear demand specification, equilibrium prices are not available in closed form. However, there exists a close mapping between the equilibrium impact of counterfactual experiments in KY (2019) and equation (1.6).

COROLLARY 1 *Let* $\log P = \mathbf{g}(\mathbf{A}, \mathbf{X}, \beta, \epsilon)$ *denote the equilibrium vector of log prices as in the demand system by KY(2019, equation 22). Any counterfactual experiment involves a change in assets under management* $\Delta\mathbf{A}$*, characteristics* $\Delta\mathbf{X}$*, demand coefficients* $\Delta\beta$ *or latent demand* $\Delta\epsilon$*, and can be represented as a change in aggregate dollar demand* $\Delta D \in \mathbb{R}^N$*. A first order approximation of the equilibrium price impact of counterfactual experiments is given by equation (1.6).*

See Appendix A for a proof. The demand-based approach by KY (2019) has been used in various settings to assess the impact of counterfactual experiments on equilibrium prices. Corollary 1 states, that the counterfactual experiments can in general be restated in the form of an aggregate demand shock ΔD: Koijen et al. (2022) assess the impact of a shift towards passive management by redistributing assets from active to passive funds. They also compute the equilibrium price effects of a stronger demand for green stocks. Han et al. (2021) evaluate the impact of mutual fund risk shifting on the beta anomaly. They redistribute underperforming funds' assets towards all other investors and compute the equilibrium price change of high beta stocks. Bretscher et al. (2020) conduct a battery of counterfactuals for the corporate bond market by changing the wealth distribution across investors A^i and bond characteristics

[3]Specifically, they derive $w^i(n) = \frac{Q^i(n)P(n)}{A^i} = e^{\beta_0^i P(n) + X(n)\beta^i + \epsilon^i}$ from mean-variance portfolio choice under the assumption of a factor structure in the covariance matrix in returns and that both expected returns and factor loadings are linear in X.

X. Jiang et al. (2020) decompose the US net foreign asset position by iteratively changing its underlying determinants and computing the counterfactual equilibrium.

The counterfactual price change is - at least to a first order - driven by $\mathscr{M}$. The quantitative implications of all of the above-mentioned counterfactuals almost entirely depend on how elastically investors respond to exogenous price changes. If investors are price inelastic (small ζ^i) then $\mathscr{M}$ is large and prices have to move a lot to accommodate the exogenous demand shock ΔD. KY (2019) estimate, that the diagonal elements of $\mathscr{M}$ range between 2 and 4, while the off-diagonal elements are close to 0. Thus, a counterfactual experiment that entails a \$1 demand shock for stock n raises the price of n by approximately \$2-4.

Lastly, note that in this simple framework the multiplier $\mathscr{M}$ is exogenous and not affected by the demand shock ΔD. Haddad et al. (2021), suggest that $\mathscr{M}$ is potentially endogenous as investors' choose their elasticity of demand in response to the trading behaviour of others. It is however likely, that changing an institution's overall trading aggressiveness is inhibited by e.g. agency frictions and does not occur instantaneously. In this light, equation 1.6 should be viewed as a *short-term* approximation of the price impact of counterfactual experiments.

1.3 Estimating Elasticities from Changes in Holdings

Investors' price elasticities are key in order to understand the role of demand shocks in financial markets. Before diving into the elasticity estimation it is worth noting that there exists extensive empirical evidence empirical evidence from index inclusions, mutual fund flows, index reconstitutions and dividend reinvestments suggesting investors have less elastic demand curves than what is implied by many frictionless neoclassical models.[4] Virtually all of the reduced form evidence for *inelastic* demand examines the relationship between *changes* in the price and *changes* in demand using suitable instruments to account for their joint endogeneity. In this spirit, I propose estimating investor-specific elasticities using *changes* in their quarterly reported share holdings as a response to non-fundamental *changes* in prices.

A Data and Variable Construction

In the US, institutional investment managers who have discretion over \$100M or more in designated 13F securities, must report their respective share holdings $S_t^i(n)$ via quarterly SEC 13F filings. I obtain institution-level holdings from 2000 to 2021 from Thomson's Institutional Holdings Database (s34 file). Monthly data on mutual funds' net returns and total net assets, as well as other fund-specific characteristics are obtained from the CRSP survivorship-bias-free mutual fund database. For over 90% of all mutual funds, CRSP provides holdings at a higher frequency than Thomson's Quarterly Mutual Fund Holdings Database (s12 file). I construct

[4] See the calibrations in Petajisto (2009) and Gabaix and Koijen (2021)

mutual fund portfolios using the both databases and opt for CRSP holdings when moving to a higher frequency. The data on quarterly stock holdings are subsequently merged with price and fundamentals data from CRSP and Compustat.[5] Institution-level and mutual fund portfolio weights $w_t^i(n)$ are constructed as the dollar holdings in each stock (price $P_t(n)$ times shares held $Q_t^i(n)$) divided by their assets under management A_t^i. An institution's assets under management are given by the sum of its dollar holdings. The subscripts t indicate the report date of the 13F filing. In line with the theory, I normalize by shares outstanding such that a stock's price and its market capitalization coincide and the number of shares held is equivalent to the fraction of ownership $Q_t^i(n) = \frac{S_t^i(n)}{\text{Shares Outstanding}_t(n)}$. Thus stock splits do not contaminate changes in shares held from one quarter to the next. Percentage changes in prices and portfolio weights are defined as $\Delta p_t(n) = \log P_t(n) - \log P_{t-1}(n)$ and $\Delta w_t^i(n) = \log w_t^i(n) - \log w_{t-1}^i(n)$. Identifying investor's elasticity of demand requires a measure of demand that captures how actively the institution is trading. I therefore define investor demand Δq_t^i as the change in portfolio weight that is not driven by valuation changes. Formally, $\Delta q_t^i(n) \equiv \Delta w_t^i(n) - \Delta p_t(n)$.[6]

B Estimation Specification

The most natural estimation, that directly emerges from the definition of the elasticity ζ_n^i in equation (2.5), is evaluating the percentage trade $\Delta q_t^i(n)$ as a response to a percentage change in the price $\Delta p_t(n)$

$$\Delta q_t^i(n) = \zeta^i \Delta p_t(n) + \epsilon_t^i(n) \tag{1.8}$$

where $\epsilon_t^i(n)$ captures demand shocks due to e.g. fundamental news, flows or legal constraints. A *causal* identification of elasticities requires *exogenous* variation in prices orthogonal to $\epsilon_t^i(n)$. For now, I will ignore the endogeneity problem. The next section shows how to construct valid instruments for prices. It is important to note that this estimation specification is defined in changes in order to elicit investors' *trades* as a response to changes in the price. The specification in changes is therefore conceptually in line with the reduced form evidence on inelastic demand mentioned that examines the price change following exogenous demand changes (such as index inclusions). In a conceptual deviation from the event-study evidence for inelastic demand, KY (2019) identify elasticities in a cross-sectional regression in *levels* as opposed to *changes*. Motivated by the fact that investors' portfolio weights are log-normally distributed in the cross-section, they estimate

$$\log \frac{w_t^i(n)}{w_t^i(0)} = \beta_{0,t}^i \log P_t(n) + \epsilon_t^i(n) \tag{1.9}$$

[5] See KY (2019) for details on the construction of the database.

[6] Investor demand $\Delta q_t^i(n)$ can also be viewed as the percentage change in shares held $Q_t^i(n)$ that is not driven by growth in assets. Intuitively, a fund that simply scales up its holdings proportional to inflows should not be interpreted as *trading actively*. To see this, note that $\Delta q_t^i(n) = \log\left(\frac{Q_t^i(n)P_t(n)}{A_t^i}\right) - \log\left(\frac{Q_{t-1}^i(n)P_{t-1}(n)}{A_{t-1}^i}\right) - \Delta p_t(n) = \log\frac{Q_t^i(n)}{Q_{t-1}^i(n)} - \log\frac{A_t^i}{A_{t-1}^i}$.

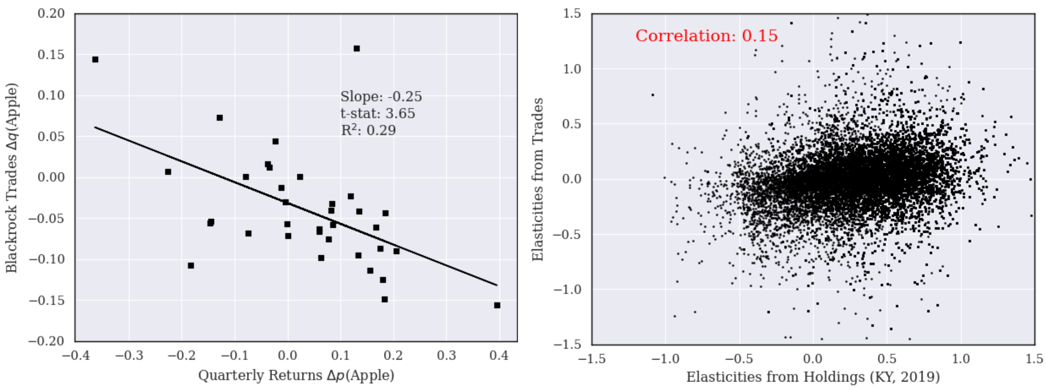
using quarterly cross-sectional regressions for each investor. Here $\epsilon_t^i(n) = \alpha_t^i + \sum_{k=1}^K \beta_{k,t}^i X_{k,t}(n) + \varepsilon_t^i(n)$ captures demand for stock-specific characteristics $X_{k,t}(n)$ and unobservable (latent) demand shocks $\varepsilon_t^i(n)$. Because shares outstanding are normalized to one, a stock's market capitalization and its price $P_t(n)$ coincide. Subtracting $\log P_t(n)$ from both sides and noting that the weight in the outside asset $w_t^i(n)$ is subsumed by the constant in a cross-sectional regression, (1.9) can be rewritten as

$$\log Q_t^i(n) = -\zeta_t^i \log P_t(n) + \epsilon_t^i(n) \tag{1.10}$$

where $\zeta_t^i = 1 - \beta_{t,0}^i$. Hence the regression using trades (1.8) is nothing but a first difference estimator of the level-specification in KY (2019). The specification in levels implies high price elasticities for investors tilting towards cheaper (controlling for fundamental value) stocks in the cross-section. The static nature of 13F investors' holdings in the time-series, which is often used as a reason for the low elasticity estimates, is not reflected in the estimated parameters. In fact, the entire demand system could be estimated using only one quarter of investor holdings and completely disregarding any time-series information. The purely the cross-sectional identification in levels also implies that elasticities are inferred as opposed to directly estimated. When conducting counterfactual experiments, the inferred elasticities do not necessarily reflect investors' true demand response as a result of an exogenous (time-series) change in the price. Lastly, institutional portfolios are to a large extent driven by (unobservable) investment mandates, which are captured by the error term $\epsilon_t^i(n)$. Any unobservable determinant of the cross-sectional portfolio choice that is correlated with prices leads to biased elasticities. E.g. a hedge fund tilting towards technology stocks will have a low ζ_t^i, simply because technology stocks have higher prices (i.e., $\text{Cov}(\epsilon_t^i(n), P_t(n)) > 0$). In order to better understand the nature of changes in holdings consider the following simple example based on true numbers. Blackrock's total number of shares in Apple in the second quarter of 2014 was $Q_{\text{2014q2}}^{\text{Blackrock}}(\text{Apple}) = 320$ Million. From the second to the third quarter of 2014 Blackrock decreased its shares to 300 Million, which implies $\Delta q_{\text{2014q2}}^{\text{Blackrock}}(\text{Apple}) \approx -6\%$. KY (2019) use the holdings of one specific quarter, run a cross-sectional regression onto prices as in (1.10) and obtain $\zeta^{\text{Blackrock}} = 0$. The level-estimate thus implies that Blackrock is perfectly inelastic with respect to Apple and all other stocks. Panel (a) of Figure 1.1 plots Blackrock's actual quarterly trades in Apple from 2010 to 2020 and the corresponding quarterly returns. The plot suggests that Blackrock's demand for Apple is not perfectly inelastic with respect to Apple's price. In fact, there is a significant negative relationship between Blackrock's demand for Apple and Apple's price. Panel (b) of Figure 1.1 compares the estimates of ζ^i for the specification in levels (Q_t^i) and changes ($\Delta q_t^i(n)$) for all investor-quarter pairs. The estimates across the two specifications are uncorrelated.

Figure 1.1: **Elasticity Estimates: Trades Δq_t^i vs Holdings Q_t^i.**
Panel (a) of the figure plots Blackrock's quarterly trades $\Delta q_t^{\text{Blackrock}}(\text{Apple})$ and Apple's quarterly returns $\Delta p_t(\text{Apple})$ from 2010 to 2020. Quarterly trades are defined in excess of asset growth as $\Delta q_t^i(n) = \log\left(Q_t^i(n)/Q_{t-1}^i(n)\right) - \log\left(A_t^i/A_{t-1}^i\right)$. The results remain unchanged if asset growth $\log\left(A_t^i/A_{t-1}^i\right)$ is computed without the valuation gain of Apple itself. Panel (b) compares the elasticity estimates for all investor-quarter pairs for the estimation in levels and changes. The x-axis plots the estimates in levels using holdings Q_t^i as in KY (2019): $\log Q_t^i(n) = -\zeta^i \log P_t(n) + X_t(n)\beta^i + \varepsilon_t^i(n)$. The y-axis plots the estimates in changes Δq_t^i as proposed in the current paper: $\Delta q_t^i(n) = \zeta^i \Delta p_t(n) + X_t(n)\beta^i + \varepsilon_t^i(n)$. The stock-specific controls $X_t(n)$ are book equity, market beta, profitability, investment and dividends-to-book equity. They do not change substantially over quarters and are hence kept in levels. The results remain unchanged if lagged controls or $X_{t-1}(n)$ or changes in the controls $\Delta X_t(n)$ are added. Both specifications are estimated without the use of instruments for endogenous prices $\log P_t(n)$ and endogenous price changes $\Delta p_t(n)$.

(a) Blackrock trading Apple (b) Trades vs Holdings



Note, that the primitive estimates in Figure 1.1 do not use instruments to address the joint endogeneity of demand (changes) and prices (changes). Therefore, the slope coefficients are not measuring the causal elasticity. Figure 1.1 should hence be viewed as tentative evidence for potentially omitted variables driving both unobservable portfolio tilts and prices. Thus the cross-sectional estimation in levels may not adequately capture investors *true* response to fluctuations in prices. Even though the original framework by KY (2019) has already been extended on multiple fronts (see e.g. Koijen et al. (2022), Haddad et al. (2021)), the cross-sectional identification in levels has been widely adopted and has remained unquestioned.

C Simulation

The different estimates from the specification in changes versus levels point towards an omitted variable bias due unobservable portfolio tilts (e.g. investment mandates) that are correlated with prices. In order to better understand the potential bias and how it may be eliminated using changes in holdings, consider the following simulation: Investor s enters the market in 2018 with 1 Billion dollars under management and equally distributes her assets across the 500 largest stocks in the US, i.e. she starts with an equal-weighted portfolio. The vector of purchased shares (as reported in her 13F filing) is given by $\theta(n) = 1\text{Billion} * \frac{1}{500 * P_{2018}(n)}$ where $P_{2018}(n)$ are the market prices as of 2018. She never trades and therefore holds the shares purchased in 2018 until the end of the sample. Her portfolio can be described as

$$\log Q_t^s(n) = \log\theta(n) + \varepsilon_t(n). \tag{1.11}$$

where $\varepsilon_t(n)$ are iid liquidity trades or reporting errors in the holdings data. $\theta(n)$ can be interpreted as an unobservable investment mandate. As she does not trade, her demand is perfectly inelastic (i.e., $\zeta_t^s = 0$). I compute $\theta(n)$ as of 2018 and then simulate her portfolio path until 2021.[7]. Elasticity estimates in levels will be unbiased as long as $\theta(n)$ and $P_t(n)$ are cross-sectionally uncorrelated. Here, $\theta(n)$ is negatively correlated to the market prices, which should lead to an *upward bias* in the elasticity estimates.[8] In other words, the estimation in levels will produce elasticity estimates that are too high. In first differences however, the unobservable investment mandate (and hence the omitted variable bias) is removed as $\Delta q_t^s = \Delta\varepsilon_t(n)$. Table 1.1 reports the estimated coefficients from the simulation. The first column reports the investor's true elasticity, which is 0. The estimation in levels produces a strong and significant upward bias in the elasticity estimates. Estimating elasticities in first differences (i.e. using trades) eliminates the bias by removing the influence of the latent investment mandate.

Table 1.1: **Elasticity Bias Simulation.**
The table reports the estimated elasticities from simulated portfolio data. I simulate the portfolio of an investor that enters the market in 2018 with an equal weighted portfolio across the largest 500 stocks and (up to iid liquidity trades $\varepsilon_t(n)$) never trades thereafter. Thus her portfolio is given by $\log Q_t^s(n) = \log\theta(n) + \varepsilon_t(n)$ where $\theta(n) = \frac{1\text{Billion}}{500 * P_{2018}(n)}$ and $\varepsilon_t(n) \sim \mathcal{N}(\mu_Q, \sigma^2_{\log Q_{2018}})$ where μ_Q and σ_Q^2 are the cross-sectional sample mean and standard deviation of $\log Q_{2018}^s$. I simulate her portfolio from 2018 until 2021 and estimate demand every quarter. I report the average coefficient over time along with its standard deviation.

		Estimates ζ^i	
	True ζ^s	Levels ($\log Q_t^s$)	Changes (Δq_t^s)
Average	0	0.92***	0.01
Std. Error	-	(0.016)	(0.024)

The simple simulation emphasizes the importance of accounting for unobservable investment mandates or portfolio tilts. The direction of the bias is driven by the correlation of the mandate with market prices. E.g. if the investor cross-sectionally tilts towards tech stocks (which are more expensive controlling for fundamental value) the elasticity estimates from levels will be biased downward. Constructing exogenous variation in prices potentially mitigates these

[7] I use $\varepsilon_t(n) \sim \mathcal{N}_t(\mu_Q, \sigma^2_{\log Q_{2018}})$ where μ_Q and σ_Q^2 are the cross-sectional sample mean and standard deviation of $\log Q_{2018}^s$.

[8] More formally, $\log\theta(n) = \log(\frac{1\text{Billion}}{500}) - \log P_{2018}(n)$ is negatively related to the market prices of 2018. Because the cross-section of prices in subsequent years is positively correlated to the prices of 2018, $\text{Cov}_t(\log\theta(n), \log P_t(n)) < 0$. In a univariate regression of $\log Q_t^s(n)$ onto $\log P_t(n)$ the bias in the estimate is given by $\frac{\text{Cov}_t(\log\theta(n), \log P_t(n))}{\text{Var}_t(\log P_t(n))}$. The point estimate is therefore biased downward which implies an overestimation of the elasticity.

concerns. The next section addresses this issue in more detail and provides an instrument for $\Delta p_t(n)$ to cleanly identify ζ^i using how investors trade as opposed to what they hold.

1.4 Causal Identification of Elasticities

In any economic context, correctly identifying an agent's causal response to the price requires exogenous variation in prices that is unrelated to the unobservable (latent) drivers of the action. E.g. identifying causal responses of managerial decisions as a response to the stock price requires exogenous variation in prices unrelated to the (in)direct effects of that decision on the price itself. The most commonly used instruments in the literature are index reconstitutions and mutual fund flow-induced price pressure. KY (2019) estimate elasticities in levels and hence construct instruments for the cross-section of prices as opposed to price changes. To this end they use exogenous cross-sectional variation in holdings via investment mandates as opposed to exogenous variation in changes in holdings. More specifically, they compute the counterfactual equilibrium prices if all investors held equal weighted portfolio given their investment universe. This instrument varies little over time and is purely aimed at cross-sectional identification in levels. Identifying how investors change their holdings, i.e., trade, as a response to changes in the prices, requires instruments for the cross-section of price changes, i.e., returns. A key advantage of estimating elasticities via trades is that essentially all of the proposed instruments from the event-study literature can be employed. In this paper, I opt for an instrument based on mutual-fund flows because it provides exogenous demand-shocks for stocks at all times. I leave the elasticity estimation using alternative instruments to future research.

A Hypothetical Trades due to Exogenous Flows

Starting with Edmans et al. (2012), flow-induced demand has been a commonly used instrument to identify causal effects in corporate finance. Formally, flow-induced demand pressure on stock n is defined as $f_{t+1}(n) = \sum_{i=1}^{I} Q_t^i(n) \frac{F_{t+1}^i}{A_t^i}$ where F_{t+1}^i are dollar flows to fund i between t and $t+1$. Thus $f_{t+1}(n)$ is a measure of all funds' hypothetical purchases and sales in n if they simply scaled their lagged holdings $Q_t^i(n)$ proportionally to relative flows $f_{t+1}^i = \frac{F_{t+1}^i}{A_t^i}$. Under the strong assumption that $f_{t+1}(n)$ is orthogonal to all other (fundamental or non-fundamental) drivers of returns one could potentially identify $\mathcal{M}$ by directly regressing returns onto $f_{t+1}(n)$. Even if one had access to purely exogenous demand shocks from unique events (as in Ben-David et al. (2020) or Han et al. (2021)), identifying $\mathcal{M}$ from a simple regression of returns onto demand is difficult as $\mathcal{M}$ is an $N \times N$ matrix and we only have NT return observations. Gabaix and Koijen (2021) investigate the single risky asset case, $\mathcal{M}$ for the aggregate stock market. They obtain flows for all sectors from the Flow of Funds data and extract idiosyncratic (non-fundamental) flows using granular instrumental variables.[9] The demand system approach takes a different route by estimating investor-specific elasticities from holdings data and - imposing market clearing - deriving a structural estimate of the complete $N \times N$ multiplier matrix. Note, that estimating price elasticities of demand is subject to essentially the same endogeneity concerns that contaminate regressions of prices onto flows. Under a successful identification, however, the structural approach gives rich insights on the underlying investor-specific determinants of the flow multiplier.

Wardlaw (2020) rightfully points towards a mechanical correlation of returns and flow-induced demand if $f_{t+1}(n)$ is scaled by contemporaneous trading volume. However, even without scaling by volume, $f_{t+1}(n)$ only provides exogenous (non-fundamental) variation in demand insofar as fund-specific flows are not driven by the fundamentals of the funds' underlying assets. If fund flows are correlated with the funds' respective portfolio tilts then $f_{t+1}(n)$ is inherently endogenous. For example deteriorating earnings of technology firms may cause flows out of mutual funds with a technology tilt. Another intuitive example are flows into ESG funds, which tilt towards or exclude unsustainable stocks. New regulation on carbon emissions may boost both the fundamental value of sustainable stocks as well as flows from conventional towards sustainable funds. On top of the potential endogeneity of flows and fundamentals, it is also possible that investors chase fund returns at a higher frequency than the observable flows (see e.g. Schmickler (2020)). If poor stock returns, which worsen mutual fund performance, cause outflows within the same month then this results in a correlation between flow-induced mutual fund demand and contemporaneous stock returns at a monthly frequency that is not purely exogenous. To address the above-mentioned concerns, note that one can decompose fund flows f^i_{t+1} into an endogenous component driven by portfolio characteristics, and an exogenous component $f^{\perp,i}_{t+1}$. Exogenous flows are then extracted using cross-sectional regressions

$$\forall t : f^i_{t+1} = C^i_t \beta_t + f^{\perp,i}_{t+1} \tag{1.12}$$

where C^i_t is a vector of portfolio characteristics. Portfolio characteristics are constructed using fund-level characteristic scores as in Lettau et al. (2018), which are portfolio-weighted averages of the stock characteristics within a fund's portfolio.[10] For every fund, I compute scores for greenness, value, size, momentum, profitability, investment and idiosyncratic volatility. As an additional portfolio characteristic, I also include fund's current return to control for contemporaneous return chasing behaviour. Exogenous flows are given by the residuals from the above regression. If households would only consider variables unrelated to funds' portfolio tilts and contemporaneous performance in their mutual fund allocation, $f^{\perp,i}_{t+1}$ and f^i_{t+1} would coincide. Using $f^{i,\perp}_t$ one can then construct hypothetical mutual fund trades. Recall that hypothetical trades are the sales and purchases of funds if they scale their

[9]They construct $f_{t+1}(n)$ by using investors' equity shares as Q^i_t, total fund inflow as F^i_{t+1}, and total assets (equity and bond holdings) as A^i_t.

[10]Formally,

$$C^i_t = \sum_{n=1} w^i_t(n) C_t(n) \tag{1.13}$$

where $C_t(n)$ is a stock-specific characteristic such as market-to-book ratio.

past holdings Q^i_{t+1} proportionally to flows. Exogenous flow-induced demand is then simply given by summing over all hypothetical trades:

$$f^{\perp}_{t+1}(n) = \sum_{i=1}^{I} f^{i,\perp}_t Q^i_t(n) \tag{1.14}$$

These flow-induced demand shocks are potentially orthogonal to the fundamental news driving returns $\Delta p_t(n)$ and can be used as instruments to causally identify investors' elasticity of demand. The next section investigates, whether the flow-induced demand shocks are indeed valid instruments, i.e., whether stocks with high flow-induced demand experience higher contemporaneous returns.

B Validity of the Instrument

If stocks with high flow-induced demand experience significantly higher returns in the same quarter, they may serve as valid instruments to identify investor-specific elasticities. To this end, I estimate the following panel regression using quarterly returns $\Delta p_{t+1}(n)$ and quarterly flow-induced demand $f^{\perp}_{t+1}(n)$ from 2010 to 2020,

$$\Delta p_{t+1}(n) = \theta f^{\perp}_{t+1}(n) + \epsilon_t(n). \tag{1.15}$$

where $\epsilon^i_t(n) = \sum_{k=1}^{K} X_{t,k}(n)\beta^i_k + \varepsilon^i_t(n)$. $X_{t,k}(n)$ are lagged characteristics known to predict returns including log market equity, book equity, market beta, 12-month momentum, idiosyncratic volatility, investment and profitability. Table 1.2 reports the estimates for θ under several different regression specifications. Consistent with downward-sloping demand for stocks,

Table 1.2: **Price Pressure from exogenous Mutual Fund Flows.**
The table reports the results of estimating 1.15 from 2010 to 2020 for different specifications. (1) is estimating the coefficient without controls or fixed effects. (2) and (3) are estimations with Controls and fixed effects respectively. (4) uses quarterly returns including dividends instead of percentage price changes $\Delta p_{t+1}(n)$. (5) is an estimation over the largest 500 stocks only. (6) uses as an dependent variable a dummy equal to 1 if $f^{\perp}_{t+1}(n)$ is in the largest quintile.

	Returns Δp_{t+1}					
	(1)	(2)	(3)	(4)	(5)	(6)
$f^{\perp}_{t+1}$	6.82***	2.65***	2.55***	2.53***	5.97***	0.01***
	(0.231)	(0.307)	(0.369)	(0.369)	(0.704)	(0.002)
Controls	No	Yes	Yes	Yes	Yes	Yes
Fixed Effects	No	No	Yes	Yes	Yes	Yes
Include Dividends	No	No	No	Yes	No	No
Largest 500 Stocks	No	No	No	No	Yes	No
Dummy($f^{\perp}_{t+1}$ Quintile 5)	No	No	No	No	No	Yes

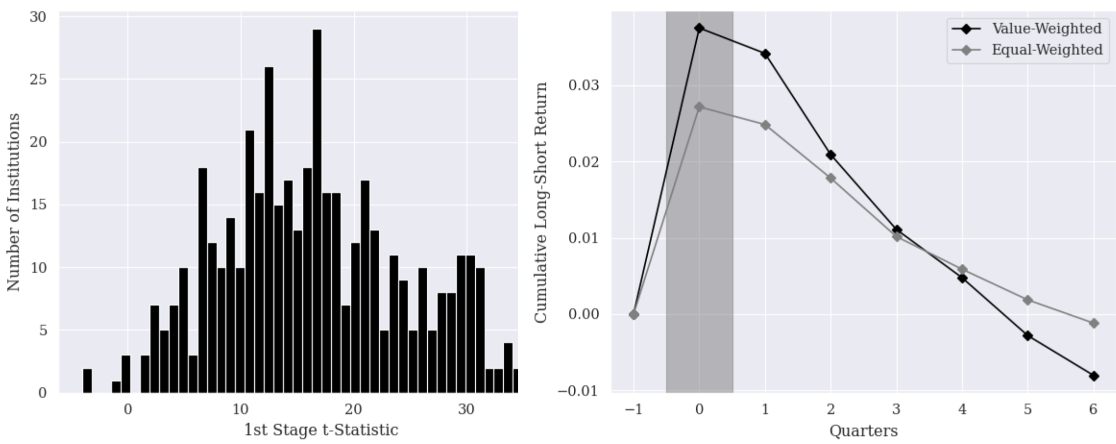
non-fundamental buying pressure due to exogenous mutual fund flows is strongly correlated with contemporaneous realized returns. The coefficient is highly statistically significant across many specifications, i.e., for different transformations of $f^{\perp}_{t+1}(n)$ and $\Delta p_{t+1}(n)$, using different controls, including fixed effects and for different subsets of stocks.

Note, that the ultimate goal is to use $f^{\perp}_{t+1}(n)$ to identify investor-specific price elasticities of demand. Thus $f^{\perp}_{t+1}(n)$ needs to be relevant for every investor, i.e., it needs to be significantly related to returns over investor-specific subsets of stocks. For many investors the number of available observations is small. This is because 13F filings are reported quarterly and because many institutions hold very few stocks in the cross-section. I therefore pool investors with fewer than 1000 stock-specific trades $\Delta q^i_t(n)$ (i.e., cross-section and time-series combined) into 9 groups based on their assets under management, diversification and trading activity. Panel (a) of Figure 1.2 reports the t-statistic on θ for the investor-specific estimation of (1.15). For 95% of all investors, the minimum t-statistic is above the critical value of 4.05 (see Stock and Watson, 2005). The remaining investors are moved to the pooled estimation.

Figure 1.2: **Relevance and Exogeneity Tests**.
Panel (a) plots the first stage t-statistic on θ for the investor-specific estimation of (1.15). Panel (b) plots the cumulative returns to long-short portfolios that go long (short) the decile of stocks with the highest (lowest) flow-induced demand $f^{\perp}_{t+1}(n)$. Stocks are sorted into portfolios at quarter 0 (the event date). Value-weights are computed using the previous quarter's market capitalization.

(a) 1st Stage Weak Instrument Test

(b) Price Pressure Reversion



Lastly, note that the exogeneity of flow-induced demand as an instrument for realized returns is difficult to test empirically. However, if the price pressure is temporary one can tentatively conclude that the flow-induced demand shocks were not driven by fundamentals. To this end, I sort stocks into deciles based on $f^{\perp}_{t+1}(n)$ and construct equal and value-weighted long-short portfolios. Panel (b) of Figure 1.2 plots the cumulative returns to the portfolios. On average, the price pressure due to flow-induced demand is transitory but nevertheless takes up to 1.5 years to revert. Appendix Section C investigates the return predictability from flow-induced demand in detail. Stocks with higher flow-driven demand have significantly lower returns over the subsequent two years. A long-short portfolio exploiting the predictability loads positively on long-term reversal and negatively on short-term momentum.

C Estimating investor-specific elasticities

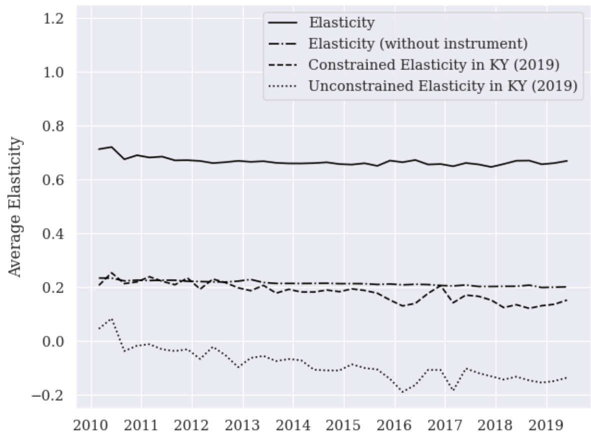
Causally identifying an investor's elasticity of demand requires variation in prices that is orthogonal to the investor's own unobservable demand shocks. As a simple illustration, consider a world in which there are only two investors, households and institutions. Then we can use the exogenous demand shocks by institutions as an instrument to identify the price elasiticity of households' demand and the exogenous demand shocks by households to identify the elasticity of institutional demand. Similarly, the flow-induced demand shocks of all institutions except j can be used as instruments to identify the elasticity of j's demand. I therefore construct $f_{t+1}^{\perp,-j}(n) = \sum_{i \in MF, i \neq j} f_t^{i,\perp} Q_t^i(n)$ to identify the elasticity of fund j.

Investor-specific elasticities are obtained in a simple two-stage least squares procedure. Let $\Delta\hat{p}_t^i$ denote the fitted value from regressing returns onto the investor-specific instrument $f_{t+1}^{\perp,-i}(n)$. The second stage regression of investor-specific trades Δq_t^i onto the investor-specific instrumented return $\Delta\hat{p}_t^i$ allows identifying their price elasticities of demand ζ^i. Formally, for every investor the two stages are given by:

$$\begin{aligned} \text{1st Stage:} \quad \Delta p_t^i(n) &= \theta f_{t+1}^{\perp,-i} + \epsilon_{t,1}^i(n) \\ \text{2nd Stage:} \quad \Delta q_t^i(n) &= -\zeta^i \Delta\hat{p}_t^i(n) + \epsilon_{t,2}^i(n) \end{aligned} \tag{1.16}$$

where $\epsilon_t^i(n) = \sum_{k=1}^K X_{t,k}(n)\beta_k^i + \varepsilon_t^i(n)$ includes the control variables log size, log book equity, profitability, investment, momentum, idiosyncratic volatility and market beta. Two things are worth noting. First, while in this specification ζ^i is a scalar, the estimation can be easily extended to the entire $N \times N$ elasticity matrix (see Section C). Second, (2.12) is a panel regression resulting in constant (as opposed to time-varying) elasticities for each investor. I opt for the pooled specification for several reasons. It captures time series information and therefore allows estimating stock-specific elasticities and cross-elasticities (section C). Cross-elasticities are given by investor i's trades in stock n as a result of a price change in m and are difficult to identify in a purely cross-sectional estimation as in KY (2019). The pooled estimation furthermore increases the number of observations for each investor and thus leads to more precisely estimated elasticities. I estimate (2.12) using quarterly trades from 2010 to 2020. In order to compare the estimated elasticities to the ones from KY (2019), I compute the AUM-weighted average coefficient across all 13F institutions. Empirically, the unconstrained estimation in levels produces negative elasticities (i.e. upward sloping demand curves) for many institutions. KY (2019) therefore impose a coefficient constraint $\zeta_t^i > 0 \quad \forall i, t$. This constraint ensures unique equilibrium prices and essentially caps many institutions' elasticities at 0. In the unconstrained estimation in changes (2.12), on the other hand, the number investors with upward-sloping demand curves is negligible. Thus no coefficient constraint is necessary. Figure 1.3 plots the average AUM-weighted elasticities from (2.12) and compares them with the estimates of KY (2019).

Figure 1.3: **Average Institutional Price Elasticities of Demand**.
The figure compares the AUM-weighted average elasticity over time for the estimation in changes with and without instruments to the elasticities in KY (2019). Elasticities are averaged over 13F institutions only, i.e., the household sector is excluded. KY (2019) add the coefficient constraint $\zeta_t^i > 0$ for all investor-quarter pairs. The dotted line at the bottom of the figure reports the average elasticities obtained from their specification without the coefficient constraint.



The top line shows that the average elasticity of demand across all institutions is around 0.7, which is in line with the event-study evidence from index inclusions. Importantly, it is over 4 times larger than the constrained elasticity of demand in KY (2019) despite using the same data. This suggests, that institutional demand is considerably more price elastic than what one would obtain using cross-sectional regressions in levels. However, demand is still far more inelastic than implied by neoclassical frameworks. The line at the bottom of Figure 1.3 reports the unconstrained average institutional elasticity in KY (2019). The fact that unconstrained institutional demand in the level-specification is upward sloping is worrying as downward-sloping aggregate demand is a necessary condition for positive equilibrium prices. All in all, Figure 1.3 and 1.1 suggest that identifying elasticities using investors' actual trades Δq_t^i as opposed to their cross-sectional holdings leads to substantially different results.

1.5 Applications and Extensions

A Integrating Alternative Elasticities in KY (2019)

The estimated elasticities using investors' trades $\Delta q_t^i(n)$ can be easily integrated in the existing demand system in levels. Restating the original demand specification by KY (2019) in terms of elasticities ζ^i yields:

$$\log \frac{w_t^i(n)}{w_t^i(0)} = (1-\zeta^i)\log P_t(n) + \sum_{k=1}^{K} \beta_{k,t}^i X_{k,t}(n) + \epsilon_t^i(n) \tag{1.17}$$

One can simply plug in the investor-specific elasticities ζ^i estimated from trades (or any alternative data source) as given and estimate the remaining coefficients including latent demand via

$$\log \frac{w_t^i(n)}{w_t^i(0)} - (1-\zeta^i)\log P_t(n) = \sum_{k=1}^{K} \beta_{k,t}^i X_{k,t}(n) + \epsilon_t^i(n) \tag{1.18}$$

Having estimated the coefficients on characteristics and latent demand the demand system can then be solved as prescribed in KY (2019). In this sense, the demand system approach can be flexibly combined with parameter estimates other (reduced-form) studies: One could take different elasticity estimates from event-studies and use them instead of "in-house" elasticities obtained from the original demand system. Using index reconstitutions, Pavlova and Sikorskaya (2022) estimate the elasticity for the household sector to be 2.26. One could hence re-estimate the demand system taking $\zeta^{\text{Households}} = 2.26$ as given. This line of thought can also be applied to the other exogenous characteristics $X_{k,t}(n)$. For example one can separately identify mutual funds' demand for sustainability using changes in ESG-scores induced by changes in the providers' ratings methodology.

B Estimates at a monthly frequency with alternative data

A known issue of using 13F filings to estimate investor-specific demand parameters is the aggregated nature of the reported holdings. Holdings are reported at the institution-level (e.g. Vanguard) as opposed to the fund level (e.g. Vanguard Selected Value Fund). Estimating elasticities at an aggregated level is problematic, as the sum of all funds holds the market which is perfectly inelastic by construction.[11] Another common criticism of identifying elasticities from 13F filings is the availability of holdings at an extremely low frequency (i.e., quarterly). It is important to note, however, that the cross-sectional identification in levels in KY (2019) is not affected by the frequency of available holdings. By running cross-sectional regressions at each point in time the estimation essentially ignores the time-series component. Even if 13F holdings for all investors were available at a daily frequency, the estimates would still represent cross-sectional snapshots of the funds' holdings as opposed to how they truly trade. The estimation in changes, on the other hand, allows capturing investors' price elasticities at a higher frequency by defining $\Delta q_t^i(n)$ over the horizon at which the holdings data is available.

For over 50% of all mutual funds in the CRSP database, holdings are available at a monthly frequency. This subsample of investors is well-suited to assess whether the elasticity estimates depend on i) the level of aggregation and ii) the estimation horizon. I first estimate the

[11]The market portfolio is simply given by the shares outstanding of each stock. Abstracting from issuances and buybacks, the market portfolio is hence passive and does not respond to changes in the price.

elasticities for the level specification of KY (2019) using monthly mutual fund holdings from CRSP.[12] Panel a) of Figure 1.4 benchmarks the estimates using monthly disaggregated mutual fund holdings against the quarterly estimates from 13F filings. Despite the different data sources, the elasticity estimates for mutual funds are almost identical and fluctuate around 0.1. This result however, is a simple artifact of the cross-sectional portfolio of mutual funds, which is close to the market portfolio. One should hence expect the estimates to remain unaffected by the estimation horizon and level of aggregation. When looking at portfolio changes, however, the estimation horizon plays a critical role in the estimation. Panel b) of Figure 1.4 compares the elasticities estimated using monthly and quarterly changes in demand. The figure shows that quarterly elasticities are around 30% larger than monthly elasticities. The larger quarterly estimates suggest that investors are more price elastic in the long run. This is in line with slow-moving capital (see Duffie (2010)) due to for example agency frictions or trading costs.

Lastly, we can judge whether the estimated elasticities from changes are in line with the fund labels from CRSP. Panel (c) plots the average elasticity for all active mutual funds in the CRSP sample. Active funds are substantially more elastic than other mutual funds. While the level estimation results in a similar ranking, the specification in changes produces elasticity estimates between 4 and 7, which is an order of magnitude larger than for the estimation in levels.

C Stock-Specific Elasticities and Cross-Elasticities

The basic specification in changes (2.12) estimates a scalar coefficient for every investor, which implies that every investor's elasticities are equal across stocks ($\zeta_n^i = \zeta^i$) and that cross-elasticities are zero ($\zeta_{n,m}^i = 0$). However, there are good reasons as to why investor-specific elasticities may differ across stocks and why cross-elasticities are non-zero. First, the simple mean-variance portfolio benchmark implies an elasticity matrix that is proportional to the inverse of the covariance matrix. This implies higher elasticities for less risky stocks and non-trivial cross-elasticities as covariances across stocks are non-zero. Second, because of both indexing and benchmarking (via e.g. optimised sampling), tracking error concerns are more pronounced for larger, less volatile stocks: Deviating from market weights for large stocks implies potentially large tracking errors. Furthermore, if an investor tilts towards the market portfolio, then active trading in the largest stocks implies a high portfolio turnover. Investment constraints can also limit the substitutability (i.e., cross-elasticity) between different stocks. When an investor with an ESG mandate faces an overpriced green stock, she is more likely to substitute towards another green stock than towards a brown stock. The specification in changes allows for a flexible incorporation of stock-specific and cross-elasticities that can be easily adjusted depending on the nature of the available holdings data. Ideally we would like

[12] Although the instrument in KY (2019) is constructed at a quarterly frequency, it can be applied directly to monthly data because it produces exogenous variation in the *cross-section*. The instrument remains highly relevant using monthly cross-sections instead of quarterly cross-sections.

to estimate the following 2nd stage regression:

$$\Delta q_t^i(n) = \sum_{m=1}^{N} \zeta_{n,m}^i \Delta \hat{p}_t(m) + \epsilon_t^i(n) \tag{1.19}$$

For quarterly holdings from 13F filings the time series observations for each investor is extremely small. This makes it difficult to estimate each element of the $N \times N$ elasticities matrix individually. However, the issue can be circumvented by parameterizing $\zeta_{n,m}^i$ as a function of observable stock-specific characteristics. The problem then reduces to estimating interaction effects as in Kelly et al. (2019). For example, one could model

$$\zeta_{n,m}^i = \begin{cases} X_t(n)\zeta_X^i & n = m \\ Y_t(n,m)\zeta_Y^i & n \neq m \end{cases} \tag{1.20}$$

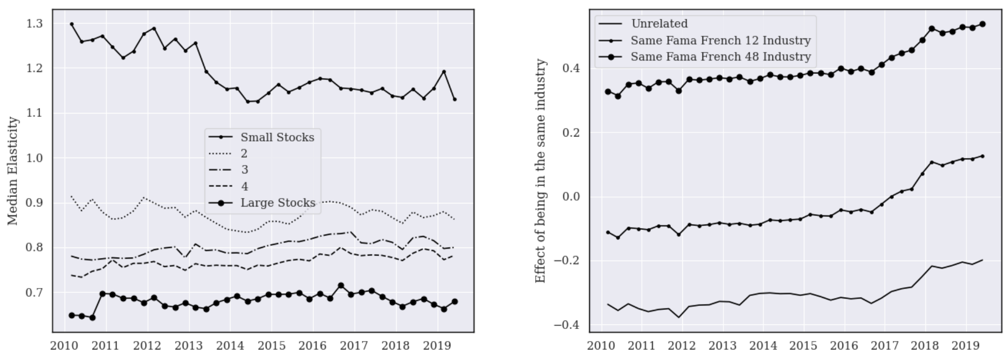
where X_t is the set of characteristics (including a constant) that describe stock-specific elasticities, such as size, market risk or idiosyncratic risk. Haddad et al. (2021) propose a similar parameterization for the level specification in KY (2019). Y_t is a set of characteristics (including a constant) describing stock-specific cross-elasticities. Importantly Y_t must encapsulate information of both n and m, as it should capture their substitutability. If n and m are very similar in terms of their characteristics, their substitutability in i's portfolio ($-\zeta_{n,m}^i$) should be greater. In essence, the similarity measure Y_t is a kernel that requires a feature map between the characteristics of n and m. In general, any kernel measuring the (dis)similarity of n and m (e.g. the $\mathscr{L}^2$ distance in the characteristics space) could be used. Here, I propose a dummy kernel equal to 1 if n and m are in the same industry. This implies that cross-elasticities can be estimated as coefficients to industry portfolios. To see this let $Y_t(n,m) = \frac{1}{|I(n)|-1} \mathbb{1}_{I(n)=I(m)}$ where $I(n)$ encodes the industry of a stock and $|I(n)|$ measures the number of stocks that are in n's industry. Plugging the parameterization in (1.19) yields

$$\Delta q_t^i(n) = \zeta_X^i \Delta \hat{p}_t(n) X_t(n) + \zeta_Y^i \Delta \hat{p}_t^{I(n)} + \epsilon_t^i(n) \tag{1.21}$$

where $\Delta p_t^{I(n)} = \frac{1}{|I(n)|-1} \sum_{m \in I(n)} \Delta \hat{p}_t(m)$ is the average return across all other stocks in n's industry. Figure 1.5 plots the stock-specific and cross-elasticity estimates averaged across institutions. I am parameterizing stock-specific elasticities as a linear function of lagged market equity, book equity, idiosyncratic volatility and market beta. For the cross-elasticities I choose dummy variables equal to 1 if m is in n's industry according to Fama French's 12 and 48 industry definitions. More elaborate parameterizations yield cross-elasticities at more granular levels. E.g. one could measure the difference in Co2-emissions between m and n in order to model the substitution between green and brown firms.

Figure 1.5: **Stock-specific and cross-elasticities**.
Panel (a) plots the median stock-specific elasticity across 5 size quintiles. Size quintiles are computed using market equity and the reported elasticity coefficients are ownership-weighted across institutions. Formally, the aum-weighted average stock-specific elasticities are computed as $z\bar{e}ta_n = \sum_i Q_t^i(n)\zeta_n^i$. Stock-specific elasticities are parameterized as a function of lagged log market equity, log book equity, market beta and idiosynratic risk. Panel (b) plots the ownership-weighted cross-elasticities $-\bar{\zeta}_{n,m} = -\sum_i Q_t^i(n)\zeta_n^i m$. The bottom line reports the cross-elasticity when n and m are neither in the same Fama French 12 industries nor in the same Fama French 48 industries. The line above reports the cross-elasticity when n and m are in the same Fama French 12 industry, but not in the same Fama French 48 industry. The top line reports the cross-elasticity when n and m are in the same Fama French 48 industry.

(a) Stock-Specific Elasticities (b) Cross-Elasticities



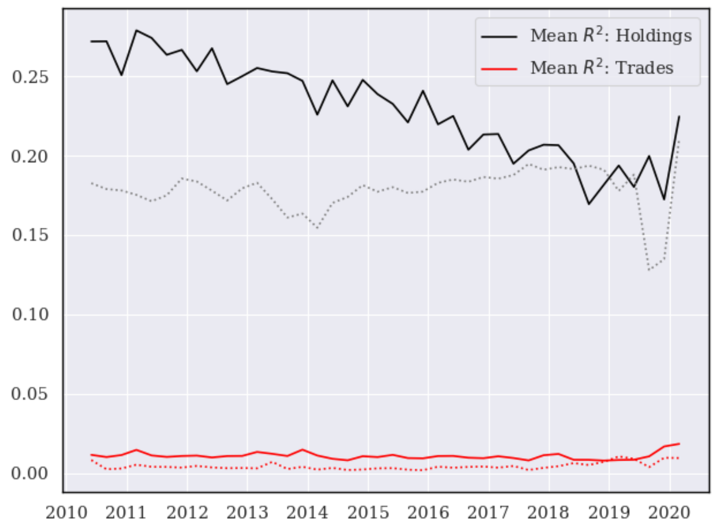
Panel (a) plots the median stock-specific elasticity for different size quintiles. In line with tracking error concerns, trading costs or turnover constraints elasticities are lower for larger stocks. Institutions are over twice as elastic with respect to the smallest quintile. The estimated elasticities for small stocks imply an average price impact of about $1/1.2 = 0.83$. Panel (b) plots the median cross-elasticities for stocks in the same Fama French 12 and 48 industry. Cross-elasticities are higher for stocks that are in the same industry. Common membership in a more narrrowly defined industry implies a higher cross-elasticity. A potential explanation is that investors like to avoid concentrated risk exposures in distinct industries. The more similar two stocks n and m in terms of their industry, the more likely they serve as substitutes and the greater the purchasing of m when the price of n increases by 1 percent. Depending on the research question, accounting for stock-specific and cross-elasticities can have strong equilibrium pricing implications. For example central bank purchases may have a greater effect on yields for bonds with a low elasticity. Alternatively, the effect of impact investors on the relative cost of capital between green and brown firms may be small when the purchased stocks have a high cross-elasticity with brown stocks.

D The Role of Latent Demand in Demand-Based Asset Pricing

The estimation in changes may not only be useful for estimating investors' elasticity of demand with respect to the price, but also their elasticity with respect to other characteristics such as ESG-scores. A potential problem with estimating characteristics-based demand in levels is that what investors *hold* cross-sectionally may not be representative of how they *trade*. A technology fund, for example, has an inherently high preference for sustainability simply because tech stocks tend to have higher ESG-Scores. Increasing a company's ESG score in a counterfactual experiment will therefore raise its stock price as a result of the incorrectly inferred demand increase by technology funds. Estimating demand for characteristics in changes alleviates many of these concerns. The issue of estimating demand via investors' *holdings* as opposed to their *trades* also extends to the role of latent demand in demand-based asset pricing. One of the most striking and perhaps disappointing findings of KY (2019) is that over 80% of the cross-sectional variation in returns is explained by changes in latent demand.[13] KY (2019) conclude that sentiment and disagreement play an important role for the cross-section of stock returns. What may not seem obvious at first sight is that the importance of latent demand in explaining the cross-section of returns is not necessarily driven by a poor cross-sectional fit of individual demand curves to the observed portfolio holdings. This (surprising) observation is directly rooted in the cross-sectional nature of the demand curves: A good cross-sectional fit of investor-specific demand curves only implies a good cross-sectional fit of *prices* and not of *returns*. E.g. it is possible that we fit individual demand curves with an R^2 of 99%, while cross-sectional return variation is entirely explained by changes in the unexplained component. In other words, if the small unexplained component (i.e., latent demand) of investors' accounts for all trades between investors, then cross-sectional variation of quarterly returns remains virtually unexplained. Figure 1.6 illustrates the difficulty in capturing *trades* as opposed to *holdings*. I plot the average AUM-weighted R^2 of explaining changes and levels of holdings with stock-specific characteristics. More specifically, I regress demand changes (levels) onto fundamentals' changes (levels). Holdings in levels are explained by fundamentals in the cross-section with an average R^2 of around 20-25%. Changes in holdings, on the other hand, are much more difficult to explain with an average R^2 of around 1%. Quarterly changes in the stock-specific variables used in KY (2019) do not seem to be the primary characteristics that investors respond to when making their quarterly trades. The figure hints at the source of KYs (2019) finding that most of the cross-sectional variation in returns is driven by latent demand. The fact that cross-sectional holdings are well explained by characteristics-based demand does *not* imply that returns are well captured by the estimated demand curves. The crux in explaining cross-sectional variation in returns via institutional demand lies in finding the drivers of trades as opposed to holdings.

[13] See KY (2019, Table 3). Because log equilibrium log prices are concave in dollar demand, variables placed first in the variance decomposition are mechanically more important in driving equilibrium prices changes. Because changes in latent demand are considered last, the 80% may in fact understate the importance of residual demand in KY (2019).

Figure 1.6: **Goodness of Fit: Holding vs. Trading.**
The figure plots the average aum-weighted R^2 of explaining changes (levels) of holdings with changes (levels) of stock-specific characteristics. Formally, the estimation in changes and levels are given by $\log Q_t^i(n) = \beta_t^i X_t(n) + \epsilon_t^i(n)$ and $\Delta q_t^i(n) = \beta_{1,t}^i \Delta X_t(n) + \beta_{2,t}^i X_t(n) + \epsilon_t^i(n)$ respectively. As explanatory characteristics $X_t(n)$, I follow KY (2019) and choose book equity, dividends to book equity, profitability, investment and market beta. Note, that the specification in changes also includes the fundamentals in levels as explanatory variables because there is little time series variation in the quarterly characteristics. I omit prices (market capitalizations) because the focus lies on the ability of exogenous characteristics in explaining portfolios. Changes Δ are quarterly and defined as $\Delta X_t = X_t - X_{t-1}$.



1.6 Conclusion

The recent literature on demand-based asset pricing infers the slope of investors' demand curves from their cross-sectional holdings. This paper shows how to estimate the elasticity of demand from 13F filings using institutions' *trades* as opposed to their *holdings*. The alternative estimation is essentially a first difference estimator of the demand curve in KY (2019). I propose using the price pressure from flow-driven mutual fund trades to identify investors' elasticity of demand. I show that institutions respond more strongly to prices than what previous estimates from the demand estimation in levels suggest. This lowers the impact of counterfactual experiments within the demand system. Furthermore, investors tend to be more elastic in the long run suggesting that the price pressure from non-fundamental demand shocks does partly revert. Overall, this paper suggest that in order to build a demand system suited to explaining *returns* as opposed to *prices* it may be worthwhile to explore characteristics that drive *trades* as opposed to *holdings*. Identifying the fundamental determinants of institutional trades is an important avenue for future research.

1.7 Proofs and Supplementary Material

A Proof of Equation (1.6)

The market clearing condition implies that $\sum_{i=1}^{I} Q^i = Q^*$ where $Q^i = f^i(P, X)$. Recall that price elasticities and cross-elasticities are defined as $\zeta_n^i = -\frac{\partial Q^i(n)}{\partial P(n)} \frac{P(n)}{Q^i(n)}$ and $\zeta_{n,m}^i = -\frac{\partial Q^i(n)}{\partial P(m)} \frac{P(m)}{Q^i(n)}$ respectively. In matrix form, the elasticity matrix is therefore given by $\zeta^i = -\text{diag}(Q^i)^{-1} \frac{\partial Q^i}{\partial P} \text{diag}(P)$. For simplicity I also define the elasticity in absolute terms (instead of percentages) as $\tilde{\zeta}^i = -\frac{\partial Q}{\partial P}$. Therefore, the elasticity in the CARA normal case with $Q^i = \frac{1}{\gamma^i} \Sigma^{-1}(\mathbb{E}[D] - P)$ is given by $\tilde{\zeta}^i = \frac{1}{\gamma^i} \Sigma^{-1}$.

We want to approximate the effects of an exogenous shock ΔX on equilibrium prices P. Differentiating both sides of the market clearing condition with respect to X yields

$$\frac{\partial Q^*}{\partial P} = \sum_{i=1}^{I} \frac{\partial Q^i}{\partial X} + \frac{\partial Q^i}{\partial P} \frac{\partial P}{\partial X} \tag{1.22}$$

Shares outstanding are normalized to 1. Therefore $\frac{\partial Q^*}{\partial P} = 0$. Rewriting (1.22) in terms of elasticities yields

$$0 = \sum_{i=1}^{I} \frac{\partial Q^i}{\partial X} - \tilde{\zeta}^i \frac{\partial P}{\partial X} \tag{1.23}$$

Now we can solve for $\frac{\partial P}{\partial X}$:

$$\frac{\partial P}{\partial X} = \left(\sum_{i=1}^{I} \tilde{\zeta}^i \right)^{-1} \sum_{i=1}^{I} \frac{\partial Q^i}{\partial X} \tag{1.24}$$

Let $\Delta Q = \sum_{i=1}^{I} \frac{\partial Q^i}{\partial X} \Delta X$ be the first order approximation to the aggregate demand shock caused by the exogenous shock ΔX. For example, positive news about the earnings of a company lead investors to update their fundamental value causing increased demand by ΔQ. A first order approximation to the equilibrium increase in prices is given by

$$\Delta P = \left(\sum_{i=1}^{I} \tilde{\zeta}^i \right)^{-1} \Delta Q \tag{1.25}$$

For example, in the CARA case the exogenous shock could be an index inclusion that reduces the supply of stocks Q^* by ΔQ. Because $\tilde{\zeta}^i = \frac{1}{\gamma^i} \Sigma^{-1}$ the equilibrium change in prices is given by $\Delta P = \gamma \Sigma \Delta Q$ where $\frac{1}{\gamma} = \sum_{i=1}^{I} \frac{1}{\gamma^i}$ is the markets effective risk aversion.

When elasticities ζ^i are defined in percentages (as in the main text), we can write the equilibrium price change as

$$\Delta P = \left(\sum_{i=1}^{I} \text{diag}(Q^i) \zeta^i \right)^{-1} \text{diag}(P) \Delta Q \tag{1.26}$$

Defining the demand shock in dollar terms $\Delta D = \text{diag}(P)\Delta Q$ yields

$$\Delta P = \left(\sum_{i=1}^{I} \text{diag}(Q^i)\zeta^i\right)^{-1} \Delta D = \mathscr{M}\Delta D \tag{1.27}$$

.

B Counterfactual Experiments in KY (2019)

Motivated by the fact that portfolio weights are log-normally distributed in the data, KY (2019) use an exponential function for $f^i(P,X)$ and model investor demand as

$$w_t^i(n) = \frac{\delta_t^i(n)}{1+\sum_n \delta_t^i(n)} \tag{1.28}$$

where $\delta_t^i(n) = \exp\left\{\beta_{t,0}^i \log P_t(n) + \sum_{k=1}^K \beta_{k,t}^i X_{k,t}(n) + \varepsilon_t^i(n)\right\}$. The sum of all investors must hold the market portfolio, thus

$$\log \sum_{i=1}^{I} A_t^i \mathbf{w}_t^i(\mathbf{p}_t) = \mathbf{p}_t. \tag{1.29}$$

In order to stay as close as possible to the notation in KY (2019), bold letters denote vectors in $\mathbb{R}^N$ and $\mathbf{p}_t = \log \mathbf{P}_t$. Due to the exponential demand specification, equilibrium log prices are not available in closed form and are a non-linear function $\mathbf{g}(\cdot)$ of all characteristics $\mathbf{X}_t \in \mathbb{R}^{N\times K}$, investors' demand coefficients $\beta_t \in \mathbb{R}^{N\times K+1}$, their assets under management $\mathbf{A}_t \in \mathbb{R}^I$ and their latent demand $\varepsilon_t \in \mathbb{R}^{N\times I}$:

$$\mathbf{p}_t = \mathbf{g}(\mathbf{A}_t, \mathbf{X}_t, \beta_t, \varepsilon_t). \tag{1.30}$$

Any counterfactual experiment changes either the wealth distribution by $\Delta\mathbf{A}_t$ (e.g. the shift towards passive management in Haddad et al. (2021)), the fundamentals by $\Delta\mathbf{X}_t$ (e.g. bond characteristics in Bretscher et al. (2020)), the demand coefficients by $\Delta\beta_t$ (e.g. higher demand for ESG characteristics in Koijen et al. (2022)) or latent demand by $\Delta\varepsilon_t$ to some alternative value. Under downward sloping demand ($\beta_{t,0}^i < 1$) the fixed point problem in (1.29) implies a unique counterfactual equilibrium price change given by $\Delta\mathbf{p}_t$. KY (2019) and all subsequent papers solve for $\Delta\mathbf{p}_t$ numerically, which makes the underlying determinants of equilibrium effects somewhat opaque. However, a first order taylor expansion produces a simple approximation for $\Delta\mathbf{p}_t$ in closed form. To this end, it is convenient to rewrite (without loss of generality) investors' demand curve in terms of a price-dependent component $\beta_{t,0}^i \log P_t(n)$ and a shock component $d_t^i(n) = \sum_{k=1}^K \beta_{k,t}^i X_{k,t}(n) + \varepsilon_t^i(n)$. The shock component captures any demand shock due to changes in characteristics, demand coefficients or latent demand. Therefore $\delta_t^i(n) = \exp\left\{\beta_{t,0}^i \log P_t(n) + d_t^i(n)\right\}$. A first order approximation to the equilibrium price change following any counterfactual experiment is hence given by

$$\Delta \mathbf{p}_t \approx \sum_{i=1}^{I} \left(\frac{\partial \mathbf{p}_t}{\partial A_t^i} \Delta A_t^i + \frac{\partial \mathbf{p}_t}{\partial d_t^i} \Delta d_t^i \right) \tag{1.31}$$

The demand shocks Δd_t^i capture demand changes due to changes in characteristics $\Delta \mathbf{X}_t$, preferences $\Delta \beta_t^i$ and latent demand $\Delta \varepsilon_t^i$. By differentiating the market clearing condition (1.29) with respect to the wealth distribution $\{A_t^i\}_{i=1}^I$ and demand shocks $\{d_t^i\}_{i=1}^I$ one can obtain closed-form solutions for the partial derivatives. I start with the assets under management. The derivative of (1.29) with respect to the assets under management of investor j, A_t^j, is given by:

$$\frac{\partial \mathbf{p}_t}{\partial A_t^j} = \mathbf{H}_t^{-1} \left(\sum_{i=1}^{I} A_t^i \frac{\partial \mathbf{w}_t^i}{\partial \mathbf{p}_t'} \frac{\partial \mathbf{p}_t}{\partial A_t^j} + \mathbf{w}_t^j \right), \tag{1.32}$$

where $\mathbf{H}_t = \sum_{i=1}^I A_t^i \text{diag}(\mathbf{w}_t^i)$. Solving for $\frac{\partial \mathbf{p}_t}{\partial A_t^j}$ yields

$$\frac{\partial \mathbf{p}_t}{\partial A_t^j} = \underbrace{\left(\mathbf{I} - \sum_{i=1}^{I} A_t^i \mathbf{H}_t^{-1} \frac{\partial \mathbf{w}_t^i}{\partial \mathbf{p}_t'} \right)}_{\mathcal{M}_t} \mathbf{H}_t^{-1} \mathbf{w}_t^j = \mathcal{M}_t \mathbf{H}_t^{-1} \mathbf{w}_t^j \tag{1.33}$$

Similarly, the derivative with respect to the other demand shocks is given by

$$\frac{\partial \mathbf{p}_t}{\partial d_t^j} = \mathcal{M}_t \mathbf{H}_t^{-1} A_t^j \frac{\partial \mathbf{w}_t^j}{\partial d_t^j} \tag{1.34}$$

The partial derivatives of $\mathbf{w}_t^j$ with respect to $\mathbf{p}_t$ and d_t^j are available in closed form. Let $\mathbf{G}_t^j = \text{diag}(\mathbf{w}_t^j)(\mathbf{I} - \mathbb{1}\mathbf{w}_t^{j\prime})$. The partial derivatives can be expressed as $\frac{\partial \mathbf{w}_t^j}{\partial \mathbf{p}_t'} = \mathbf{G}_t^j \beta_{0,t}^j$ and $\frac{\partial \mathbf{w}_t^j}{\partial d_t^{j\prime}} = \mathbf{G}_t^j$. Thus we can express (1.31) as

$$\Delta \mathbf{p}_t \approx \mathcal{M}_t \left(\sum_{i=1}^{I} \mathbf{H}_t^{-1} \mathbf{w}_t^i \Delta A_t^i + \mathbf{H}_t^{-1} \mathbf{G}_t^i \Delta d_t^i \right) \tag{1.35}$$

Note, that $\mathbf{H}_t^{-1}(\mathbf{w}_t^i A_t^i)$ is the vector of ownership shares of investor i in all stocks. Because shares outstanding are normalized to 1, this is equivalent to Q_t^i. Furthermore, the log price change is approximately equal to the percentage change, i.e., $\Delta \mathbf{p}_t \approx \text{diag}(P_t)^{-1} \Delta P_t$. Thus we can express the change in prices as the product of a dollar demand shock and the multiplier matrix

$$\Delta P_t = \mathcal{M}_t \Delta D_t \tag{1.36}$$

where

$$\Delta D_t = \sum_{i=1}^{I} \underbrace{\text{diag}(P_t) Q_t^i}_{D_t^i} \underbrace{\left(\frac{\Delta A_t^i}{A_t^i} + (\mathbf{I} - \mathbb{1}\mathbf{w}_t^{j\prime}) \Delta d_t^i \right)}_{\Delta D_t^i / D_t^i} \tag{1.37}$$

is the demand shock $\Delta D_t^i / D_t^i$ due to flows $\frac{\Delta A_t^i}{A_t^i}$ and portfolio changes $(\mathbf{I} - \mathbb{1}\mathbf{w}_t^{j\prime})\Delta d_t^i$ times the original dollar holdings $D_t^i = \text{diag}(P_t) Q_t^i$. Note, lastly, that as in KY (2019) the relationship between an investor's elasticity ζ_t^i and $\beta_{0,t}^i$ is given by

$$\zeta_t^i = \mathbf{I} - \beta_{0,t}^i \text{diag}(\mathbf{w}_t^i)^{-1} \mathbf{G}_t^i \tag{1.38}$$

See equation (14) in their paper. Plugging into the expression for $\mathcal{M}_t$ in (1.33) yields

$$\mathcal{M}_t = \sum_{i=1}^{I} \text{diag}(Q_t^i) \zeta_t^i \tag{1.39}$$

C Additional Figures and Robustness Checks

Return Predictability from Price Pressure

Flow-induced demand from mutual funds $f_{t+1}^{\perp}$ has a large and significant contemporaneous effect on prices. This price pressure tends to slowly revert over time as suggested in Figure 1.2. To further investigate the significance of the reversal, consider the following panel regression

$$r_{t+1,t+h}(n) = \theta f_{t+1}^{\perp}(n) + Controls + \epsilon_{t+1}(n) \tag{1.40}$$

where $r_{t+1,t+h}(n)$ is the return from $t+1$ to $t+h$ with $h = 12, 24, 36$ denoting the horizon in months. Thus the return excludes the contemporaneous price pressure between t and $t+1$. The control variables include stock and time fixed effects, log market equity, log book equity, profitability, investment, momentum, idiosyncratic volatility and market beta. Panel (a) of Table 1.3 reports the estimated coefficient θ for the contemporaneous return r_t and the long horizon return $r_{t+1,t+h}$. Stocks with higher flow-driven demand at time t have significantly higher returns at t and significantly lower returns over the subsequent two years. The results suggests, that the price pressure from flow-driven demand shocks is non-fundamental and is corrected over time. Panel (b) reports the performance of a long-short investment strategy that exploits the long-term reversal of flow-driven demand. The strategy goes short (long) the decile of stocks with the highest (lowest) flow-driven demand. I report the returns, alphas and factor loadings to equal and value-weighted versions of the strategy respectively. The strategy loads negatively on short-term momentum and positively on long-term reversal.

Table 1.3: **Return Predictability from Flow-Driven Demand.**
Panel (a) reports the return predictability estimates from flow-driven mutual fund demand from 2010 to 2020. I control for log market equity, log book equity, profitability, investment, momentum, idiosyncratic volatility and market beta, as well as stock and quarter fixed effects. Contemporenous returns are denoted by r_t and future returns up to h months ahead are denoted by $r_{t+1,t+h}$. The standard errors are robust to heteroskedasticity and autocorrelation. Panel (b) reports the returns to a long-short strategy that goes short (long) the decile with the highest (lowest) flow-driven demand. I report the raw long-short returns, alphas and factor loadings of the value-weighted (vw) and equal-weighted (ew) strategy. The momentum factor is based on past year returns and reversal factor is based on past 5 year returns. All factor returns are from Kenneth French's website.

(a) Long Horizon Predictability

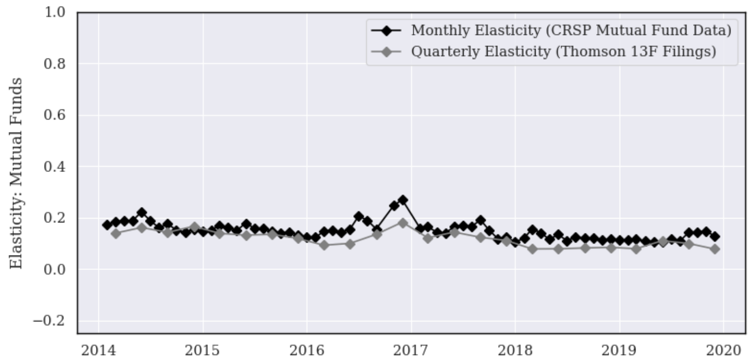
		Return Horizon $r_{t+1,t+h}$		
	r_t	h =12	h =24	h =36
$f_t^{\perp}$	0.28***	-1.39***	-0.1***	0.101
	(0.848)	(0.24)	(0.167)	(0.137)
Controls	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes

(b) Long-Short Portfolio

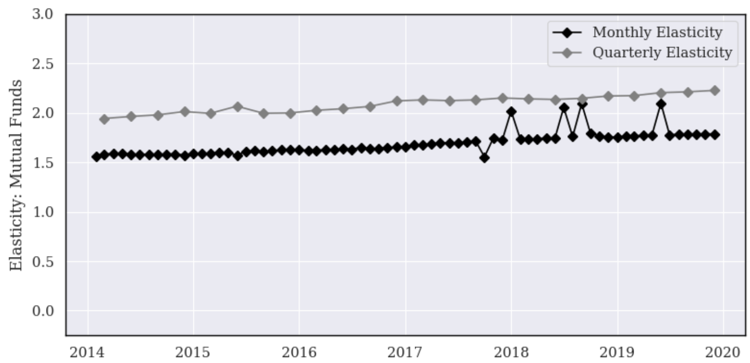
			Factor Loadings				
	Raw Return	α	β_{Mkt}	β_{HML}	β_{SMB}	β_{MOM}	β_{REV}
Long-Short Deciles (vw)	0.32	0.29*	0.05	-0.24**	0.04	-0.18***	0.21***
	(0.225)	(0.171)	(0.049)	(0.099)	(0.084)	(0.046)	(0.075)
Long-Short Deciles (ew)	0.15	0.17	-0.02	-0.43***	-0.00	-0.11***	0.35***
	(0.277)	(0.202)	(0.045)	(0.129)	(0.065)	(0.036)	(0.086)

Figure 1.4: **Mutual Fund Elasticities: Monthly versus Quarterly Frequency**.
The figure plots average aum-weighted elasticities for all active mutual funds from 2014 to 2020. Panel (a) reports estimates using the level specification of KY (2019). Quarterly estimates are obtained using quarterly 13F filings at the institution level. Panel (b) plots the estimates using the specification in changes. Quarterly and monthly estimates are obtained by constructing $\Delta q_t^i(n)$ and $\Delta p_t(n)$ as quarterly and monthly changes respectively. Panel (c) plots the estimates for all active mutual funds in the CRSP sample using the specification in changes. Active funds are defined as funds labelled by CRSP as "index enhancers".

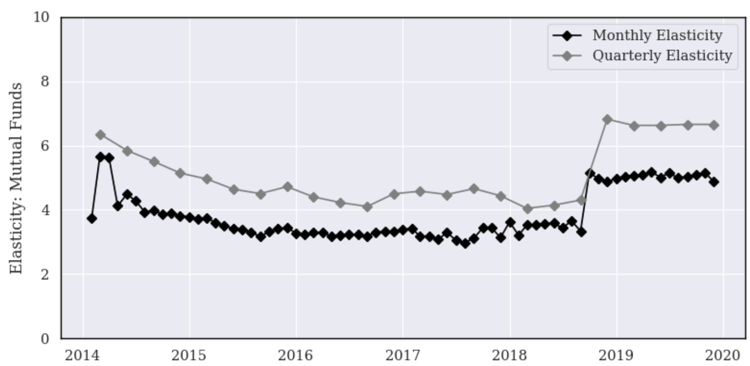
(a) Estimation in levels Q_t^i (KY (2019))



(b) Estimation in changes Δq_t^i



(c) Estimation in changes Δq_t^i: Active Mutual Funds



Chapter 2

Flow-Driven ESG Returns

Over the past decade, the sustainable investment industry has grown drastically. The high demand for sustainable investments has fueled the emergence of new funds that incorporate environmental, social, and governance (ESG) criteria into their investment decisions. Despite the enormous growth in the ESG investment industry, both the price impact and the expected returns of sustainable investing are widely debated. Academic and practitioner views on the expected returns from sustainable investments are often diametrically opposed. The pervasive theoretical view is that if investors have a preference for sustainability, the additional utility gained by investing sustainably should be offset by lower expected returns. Investors bid up the price of sustainable companies, and risk-adjusted expected returns must unambiguously be lower. In other words, investors cannot do well by doing good. Empirically, however, sustainable funds have performed well in recent years suggesting that ESG-concerned investors are in fact doing well by doing good. At the same time, the extent to which sustainable investors can impact prices is highly debated. For every buyer there is a seller. Hence, divesting from oil companies simply implies a change in ownership towards funds without a sustainability mandate. The impact of sustainable investing therefore depends on how much prices have to change in order to induce other investors to hold the divested oil shares.

This paper reconciles the price impact and realized returns of sustainable investing. I show that the high realized returns from sustainable investing are primarily driven by the price impact of flows towards sustainable funds. Flows towards ESG funds - regardless of whether they are motivated by growing ESG concerns or past fund performance - create buying pressure on the stocks that the funds overweight. This buying pressure affects prices, if the market's willingness to accommodate the demand by substituting between stocks is finite. In other words, if the aggregate demand curve for green stocks is downward-sloping, then ESG flows increase the price of green stocks. In equilibrium, the price impact of flows towards ESG funds is driven by two factors: The deviation of ESG funds from the market portfolio and the aggregate willingness to substitute between stocks (henceforth, elasticity of substitution). If the investors holding green stocks substitute elastically between stocks, then the price pressure due to ESG flows has a negligible impact. Small price changes induce investors to rebalance their portfolios by substituting away from the overpriced green stocks. On the other hand, if the holders of green assets do not aggressively rebalance their portfolios, i.e. if they are inelastic, then flows have a large price impact. I show that institutions' ability to accommodate ESG demand is limited as their elasticity of substitution between stocks is low. Thus, flow-driven trades by ESG funds have a large impact on prices, both in the cross-section of individual ESG stocks and in the time-series of ESG portfolio returns. Therefore, the *realized* returns from sustainable investing over the past decade have a large flow-driven component. The outperformance of ESG funds should hence not be interpreted as *expected* outperformance going forward.

I start by identifying a set of 551 sustainable mutual funds (henceforth ESG funds) by matching their names with a list of sustainability keywords. Using data on mutual funds' portfolio holdings, I then construct a representative ESG portfolio that pools the holdings of ESG mutual funds. The representative ESG portfolio outperformed the aggregate mutual fund portfolio with a significant 5-factor alpha of 1.51% annually. Exposure to the Green Factor by Pástor et al. (2021) does not explain the outperformance. The ESG portfolio's deviations from the aggregate mutual fund portfolio are a revealed preference measure of how sustainable a stock is (perceived to be). I define 'green' stocks as the ones overweighted by ESG funds relative to other mutual funds. Thus, in this paper 'green' refers to all dimensions of sustainability, not only environmental concerns. Quantifying the flow-driven component of the realized returns from the ESG portfolio requires a measure of total ESG flows. To this end, I propose a new measure of total capital flows into managed portfolios. The measure includes the portfolio tilts of all institutional investors and is constructed by projecting fund-specific holdings onto managed portfolios in the portfolio-weight space as opposed to the return space. Total institutional flows into the ESG portfolio amounted to \$1.3 trillion, which dwarfs the flows into ESG mutual funds of \$350 billion.[1]

In order to quantify price impact of ESG flows, I estimate a structural model that jointly matches flows and realized returns along the lines of Koijen and Yogo (2019). I use the estimation proposed by van der Beck (2022), which identifies elasticities from investors' trades, as opposed to their portfolio holdings in levels. The model allows for estimating institutions' elasticity of substitution between green and other stocks. I use demand shocks from dividend reinvestments by Schmickler and Tremacoldi-Rossi (2022) as an instrument to address the endogeneity of prices in the elasticity estimation.[2] I show that the estimates are robust to an alternative identification that uses changes in benchmarking intensities by Pavlova and Sikorskaya (2022) as an exogenous shock to supply. The estimated elasticities can be combined with ownership shares into a cross-sectional multiplier matrix. The multiplier matrix is the

[1] See Morningstar's 2021 Sustainable Funds U.S. Landscape Report.

[2] Starting with Edmans et al. (2012), flow-induced trades by mutual funds have been commonly used as exogenous demand shocks to identify causal relationships. See Wardlaw (2020) for a summary of the literature. Note, that the instrument used in this paper is immune to the Wardlaw-critique (see van der Beck (2022) for details).

cross-sectional pendant to the macro multiplier in Gabaix and Koijen (2021). I find that, ceteris paribus, a 1% demand shock for a green stock leads to a 1.11% percent increase in the price of that stock. Furthermore, cross-elasticities suggest that as money is flowing out of the fossil fuel industry, investors substitute towards green stocks, creating positive spillover effects.

I apply the model to the assess the flow-driven component in the realized returns of green stocks. The price pressure due to a $1 flow into the ESG portfolio funds is given by the product of the multiplier matrix and the deviation of the ESG portfolio from market weights. I show that every dollar flowing from the market portfolio into the representative ESG portfolio increases the aggregate value of green stocks by $0.4. I then compute the counterfactual realized returns if the total ESG flows were instead invested in the market portfolio. The price pressure from ESG flows accounts for virtually all of the outperformance of the ESG portfolio over the market portfolio in recent years. In the absence of ESG flows, the ESG portfolio would have underperformed the market with an annualized 5-factor alpha of -0.3%. This suggests that, in the absence of flow-driven price pressure, investors would have had to pay a premium for investing according to their ESG preferences. Moving to the cross-section, I show that green stocks with higher flow-driven demand had significantly higher abnormal returns. The average price impact implied from cross-sectional regressions is 1.17, which is strikingly close to the structural estimate of 1.11. Furthermore, the stocks with higher multipliers implied by the structural model indeed have a higher price impact in the cross-section, i.e. they are more affected by ESG demand.

Lastly, I provide reduced-form evidence on the price pressure of ESG demand from inclusions in the Vanguard 4Good index, as in Berk and van Binsbergen (2022). I show that not all inclusions are followed by index-tracking mutual funds. However, the inclusions with high trading volume by index-trackers are associated with significantly higher returns. The price impact implied by the index inclusions ranges from 0.87 to 1.69, which is close to the structurally estimated ESG multiplier. I then show how to decompose ESG mutual funds' portfolio additions into a mandate-driven and a fundamental component. This extends the concept of ESG index inclusions to a broader set of stocks. The mandate-driven portfolio additions across ESG mutual funds represent non-fundamental shocks to ESG demand and are significantly related to contemporaneous returns, controlling for changes in fundamentals and known return predictors. The magnitude of the estimated price impact is once again in line with the structural multiplier. Lastly, I show that there is substantial heterogeneity in the price impact of different ESG mutual funds - both in terms of magnitude and direction. Many 'sustainable' mutual funds deviate very little from S&P500 weights. Others positively affect the aggregate market capitalization of the fossil fuel industry. I outline how the model can be used to distinguish ESG funds by their *true* impact on green firms' cost of capital rather than by flamboyant fund prospectuses.

Related Literature

The empirical evidence on the *realized* returns from ESG investing over the past two decades is mixed and tends to depend on the sustainability measure, time horizon, and asset universe under investigation. Hong and Kacperczyk (2009) find that stocks in the tobacco, alcohol, and gaming industry (i.e. sin stocks) outperform other stocks. Bolton and Kacperczyk (2021a) and Bolton and Kacperczyk (2021b) find evidence for a carbon premium, implying that high-emission stocks have higher returns after controlling for known risk factors. Similarly, Hsu et al. (2020) find significant outperformance of high chemical emission stocks versus low ones.[3] Except for Hong and Kacperczyk (2009), who argue for a taste-based explanation along the lines of Fama and French (2007), these papers suggest that sustainable firms offer hedges against adverse climate events and hence require lower returns in equilibrium.

Conversely, many papers show that sustainable stocks had higher realized returns than other stocks. Edmans (2011) shows that a portfolio of firms with high employee satisfaction has a significantly positive alpha. In et al. (2020) find that an ESG portfolio, which longs low emission and shorts high emission stocks earns a significantly positive annualized alpha of 3.5–5.4%. Similarly, Görgen et al. (2020) find that from 2010 to 2017 brown (high carbon) firms performed worse than green firms on average. Hong et al. (2019) find that the risk of drought negatively predicts a country's stock returns. These papers typically propose under-reaction as a reason why sustainability is associated with a positive return premium. Pedersen et al. (2020) propose an equilibrium model with green preferences and ESG scores that are informative about stocks' risk and return. Their model shows that green stocks can have higher returns if ESG scores positively predict returns in a way that has not been appreciated by all investors. In support of the under-reaction hypothesis, Derrien et al. (2021) find that analysts downgrade their earnings forecast in response to negative ESG incidents. Glossner (2021) shows that ESG incidents predict future ESG incidents and that the stock market underestimates the adverse value effects of negative poor ESG practices. Glossner (2021) also suggests that ESG mutual funds benefit from the under-reaction to ESG news. Similarly, I argue that ESG mutual funds benefit from flow-driven price pressure on green stocks.

The theories developed in Pástor et al. (2021) and Pedersen et al. (2020) imply that the *expected* returns of green stocks should be lower than for brown stocks as investors have a taste for green assets. However, if green preferences (e.g. via climate concerns) strengthen unexpectedly over the estimation horizon, green stocks may have higher realized returns than brown firms. Alternatively, if climate risks increase unexpectedly, the hedging benefits of holding green stocks improve, which pushes up their price, resulting in higher realized returns. Thus unexpected shifts in the aggregate demand for green assets may drive a wedge between expected and realized returns. This divergence between realized and expected returns may explain the strong ambiguity in the empirical findings mentioned above. In a follow-up paper, Pastor,

[3] Other papers documenting a positive return premium on brown investments in equities, bonds, real estate, and option markets include Faccini et al. (2021), Huynh and Xia (2021), Seltzer et al. (2022), Bernstein et al. (2019), Baldauf et al. (2020), Painter (2020), Goldsmith-Pinkham et al. (2021) and Ilhan et al. (2021).

Stambaugh and Taylor (henceforth PST, 2022) regress the realized returns of a green-minus-brown (GMB) factor onto several proxies of unexpected shocks to climate concerns. They then approximate the wedge between *expected* and *realized* factor returns as the return component explained by green demand. As one particular proxy for green demand (by investors rather than consumers) they use flows to ESG mutual funds and find no significant correlation to contemporaneous GMB returns. However, ESG flows may not directly target the GMB portfolio. Instead, they flow into the aggregate ESG mutual portfolio, the returns of which are highly significantly correlated to total ESG flows. It is important to note, that this interpretation of the relationship between green demand and realized returns is slightly different from the mechanism proposed in this paper. In PST (2022), ESG demand only correlates with returns as long as it represents aggregate shifts in green preferences (or equivalently, wealth-weighted individual tastes as in Fama and French (2007)). Thus, if flows to green funds were driven by e.g. past return performance instead of growing climate concerns, prices would remain unchanged. Similarly, ESG demand shocks for individual stocks have negligible price effects as they do not change the exposure to common risk factors and have little impact on aggregate market risk.[4] Nevertheless, despite the different interpretations of the correlation between flows and returns, this paper shares the objective of measuring how the demand for green stocks affects the wedge between realized and expected returns. The joint endogeneity of prices and holdings makes identifying the causal relationship between demand shocks and realized returns extremely difficult. Simple regressions of returns onto flows are typically biased as the number of endogenous variables affecting both demand and prices are countless. Hence, this paper circumvents direct regressions of returns onto flows and instead estimates the coefficient linking flows to returns within a structural model. In a closely related paper, Berk and van Binsbergen (2022) calibrate the potential impact of ESG divestment in a frictionless CAPM world. They argue that the equilibrium price impact of sustainable investing is negligible because the high return correlation between green and brown stocks makes them strong substitutes. Therefore, inducing other investors to hold brown stocks requires little price concessions. Petajisto (2009), however, shows that the price impact implied by the CAPM greatly underestimates the estimates from the index inclusion literature. In other words, the frictionless mean-variance benchmark considerably overestimates investors' elasticity of substitution between stocks. I directly estimate demand elasticities from holdings data and show that the market's elasticity of substitution between green and brown stocks is indeed very low. Thus investors require large price concessions to accommodate the flow-driven trades by ESG funds. Using index inclusion as in Berk and van Binsbergen (2022), I find that the stocks purchased by ESG index trackers have significantly higher contemporaneous returns.

The paper also relates to the extensive literature on demand-driven price pressure. Shleifer (1986) shows that index inclusion leads to positive realized returns as a result of buying pressure by index funds. Coval and Stafford (2007), Frazzini and Lamont (2008), Edmans et al. (2012) and Lou (2012) find evidence for cross-sectional price pressure resulting from mutual funds' flow-driven trades. More recently, Parker et al. (2020) find that the rebalancing

[4] See Petajisto (2009) for a simple calibration.

of target date funds affects both the cross-section of realized returns and the aggregate stock market. Using a Morningstar ratings reform as an instrument, Ben-David et al. (2020) show that demand pressures affect the cross-section of realized style returns. Similarly, I show that mandate-driven portfolio additions by ESG funds affect individual stock returns. More closely related, Gabaix and Koijen (2021) use a structural approach to estimate the multiplier linking flows and aggregate stock market returns. They find that a $1 *unexpected* flow raises the value of the aggregate equity market by around $5. Using aggregate dividend reinvestments as an instrument, Hartzmark and Solomon (2021) find, that even *expected* uninformed flows into and out of the aggregate stock market have a price multiplier of 1.5 to 2.3. Pavlova and Sikorskaya (2022) introduce a new measure, Benchmarking Intensity, which quantifies the fraction of a stock's total market cap that is held by benchmarked investors. They show that changes in a stock's Benchmarking Intensity are an effective change in supply that is significantly related to contemporaneous returns around the Russell 1000/2000 cutoff. They find that institutional trades have a price multiplier of 1.5. As a robustness check, I use changes in Pavlova and Sikorskaya's Benchmarking Intensity as an instrument to identify investors' demand elasticities. The resulting ESG multipliers are strikingly similar to the flow-based identification. The estimated elasticities furthermore exhibit the same cross-sectional patterns.

Lastly, this paper relates to the growing literature on demand system asset pricing following the influential work by Koijen and Yogo (henceforth KY, 2019). KY (2019) present a structural model that estimates investor-specific demand curves from quarterly 13F filings and links the estimated demand coefficients to equilibrium asset prices. In a follow-up paper, Koijen et al. (2022) estimate investors' demand for environmental scores and show that long-term investors, passive funds, and banks benefit the most from growing climate concerns. Similarly, Noh and Oh (2022) regress institutional portfolio weights onto ESG-Scores and show that ESG demand predicts firm-level improvements in Co2 emissions.[5] In this paper, I refrain from explicitly estimating investors' green preferences as there are many unobserved characteristics correlated with ESG-Scores that drive demand. A technology fund, for example, has an inherently high preference for sustainability simply because tech stocks tend to have higher ESG-Scores. A valid identification of green preferences, therefore, requires exogenous variation in ESG-Scores uncorrelated with investors' unobservable investment mandates and portfolio tilts. Van der Beck (2022) proposes identifying the demand elasticities in KY (2019) from investors' trades, that is changes in their portfolios, as opposed to their cross-sectional holdings. This alleviates the concern of slow-moving unobservable variables (such as investment mandates) that drive investors' holdings in the cross-section and are correlated with prices. The estimation in changes furthermore allows identifying elasticities with existing in-

[5] A growing number of papers applies the framework by KY (2019) to estimate the impact of counterfactual experiments on equilibrium asset prices. Han et al. (2021) evaluate the impact of mutual fund risk shifting on the beta anomaly. Bretscher et al. (2020) estimate a demand system for corporate bonds. Jiang et al. (2020) use the demand system to decompose the variation in the US net foreign asset position into its underlying determinants. Van der Beck and Jaunin (2021) investigate the impact of retail traders on the equity market through the demand system approach. Haddad et al. (2021) suggest that the elasticity in KY (2019) is potentially endogenous as investors strategically update their elasticity in response to the aggregate elasticity.

struments from the reduced-form literature on price pressure.[6] This paper uses the estimation from van der Beck (2022) to identify the substitutability of green and brown stocks and links it to the realized returns from ESG investing.

The remainder of this paper is structured as follows. Section 3.2 describes the data. In Section 3.3, I construct the representative ESG portfolio. Section 3.4 briefly outlines the structural model and estimates the markets' willingness to substitute between stocks. Section 3.5 uses the model to quantify the impact of aggregate ESG flows on the time-series of ESG returns. Section 3.6 investigates stock-specific flows and the cross-section of ESG returns. Section 2.6 provides robustness tests and applications. Section 2.7 concludes.

2.1 Data and Variable Construction

A Prices and Fundamentals

Stock price data on common ordinary shares (share code 10 and 11) traded on the NYSE, AMEX and Nasdaq (exchange code 1, 2 and 3) are from CRSP. Accounting data are from Compustat. Stocks are indexed by n. Stock n's market equity as of date t is denoted by $P_{t,n}$. I normalize shares outstanding to 1, such that prices and market equity coincide. I construct the stock-specific characteristics book equity, market beta, profitability, investment, idiosyncratic volatility, turnover, momentum and industry affiliation.[7] For industry classifications, I use the Fama and French 12 industries. I furthermore construct monthly cash dividends (distribution code 1000-1399) by summing over payment dates from CRSP's daily security file. Sin stocks are defined following Hong and Kacperczyk (2009) as companies involved in the production of alcohol, tobacco and gaming. I further define controversial stocks following MSCI's exclusionary screens as companies in the biotech, firearms, oil, military and cement industry. Because a firm's sustainability is difficult to quantify and because ratings across providers often diverge strongly (see Berg et al. (2019)), I construct an objective measure using portfolio tilts of ESG mutual funds (see next section). As a robustness check, I also use Co2 emissions and ESG Scores from Refinitiv as a measure of a firm's greenness.

B Holdings and Flows

In the US, institutional investment managers who have discretion over $100M or more in designated 13F securities must report their respective holdings via quarterly SEC 13F filings. I obtain institution-level holdings from 2010 to 2021 from Thomson's Institutional Holdings

[6] E.g. Index inclusions by Shleifer (1986), mutual fund fire sales by Coval and Stafford (2007), flow-driven trades by Lou (2012), and dividend reinvestments by Schmickler and Tremacoldi-Rossi (2022) and Hartzmark and Solomon (2021).

[7] See Appendix Section D for details on the construction of these variables.

Database (s34 file). The holdings data are subsequently merged with characteristics data from CRSP and Compustat.[8] Institutions are indexed by i. I define institution i's quantity demanded $Q^i_{t,n}$ in stock n at time t as the shares held normalized by shares outstanding.[9] Institution-level and mutual fund portfolio weights $w^i_{t,n} = \frac{Q^i_{t,n} P_{t,n}}{A^i_t}$ are constructed as the dollar holdings in each stock (price times shares held) divided by their assets under management A^i_t. An institution's assets under management are given by the sum of its dollar holdings. In order to ensure market clearing, I follow KY (2019) and construct a household sector as the residual shares outstanding not held by 13F institutions.[10]

Monthly data on mutual funds' holdings, net returns and total net assets, as well as other fund-specific characteristics are obtained from the CRSP survivorship-bias-free mutual fund database. For over 90% of all mutual funds, CRSP provides holdings at a higher frequency than Thomson's Quarterly Mutual Fund Holdings Database (s12 file). I construct mutual fund portfolios using both databases and opt for CRSP holdings when moving to a higher frequency. For all mutual funds, I compute flows as $f^i_t = \frac{A^i_t - A^i_{t-1}(1+r^i_t)}{A^i_{t-1}}$ where A^i_t are the fund's total net assets and r^i_t is the monthly return between $t-1$ and t as reported on CRSP.

2.2 ESG Mutual Funds

A Identifying ESG Mutual Funds

I use fund names from CRSP's Mutual Fund Database to identify a comprehensive set of ESG mutual funds. To this end, I match fund names with a list of sustainability keywords and identify 551 ESG funds. Specifically, I define a mutual fund to be an 'ESG fund' if its name contains at least one (or any abbreviation) of a list of sustainability keywords.[11] Appendix Section A reports the largest 30 identified ESG funds as well as robustness checks to the identification of the ESG label. I then match the ESG funds with their quarterly and monthly stock holdings from both CRSP and Thomson's Mutual Fund Holdings Database (s12 file). Table 2.1 provides summary statistics on the sample of ESG funds and their aggregate portfolio. From 2010 to 2021 the average ESG fund held around 200 stocks in its portfolio. The average assets have remained relatively stable and only increased in recent years to \$630 Million. The fifth column of Table 2.1 reports the total number of ESG name changes in a given year. Out of the sample of ESG funds, 99 went from 'non-ESG' to 'ESG' by changing their name to include

[8] See KY (2019) for further details on the construction of the database.

[9] Formally $Q^i_{t,n} = \frac{\text{Shares Held}^i_{t,n}}{\text{Shares Outstanding}_{t,n}}$, where Shares Held$^i_{t,n}$ is the number of shares reported in i's 13F Filing.

[10] Furthermore, institutions with less than \$10 million under management or without any holdings in the inside and/or outside assets are attributed to the household sector, which therefore includes households, small asset managers, and other non-13F institutions.

[11] The list of sustainability keywords used is: *Environment, social, governance, green, sustainable, responsible, SRI, ESG, climate, clean, carbon, impact, fair, gender, solar, earth, renewable, screen, ethical, conscious, CSR, thematic.* See Appendix Section A for details

Table 2.1: **ESG Funds Summary Statistics**
The table reports yearly averages of quarterly metrics describing the sample of ESG funds. The first 5 columns report statistics at the ESG fund level. Excess Flows measure the average quarterly flow across all ESG funds in excess of the average quarterly flow across non-ESG funds. The number of name changes captures the total number of funds that change their name in a given year by including an ESG keyword. The last 3 columns report statistics for the aggregated portfolio of all ESG funds. Index ESG funds are ESG funds that directly track (or are based on) an index. The fraction of indexed AUM is computed as index ESG funds' total AUM relative to the total AUM of all ESG funds. Active share (Cremers and Petajisto (2009)) is computed as the deviation of the ESG portfolio $w_{t,n}^{ESG}$ from market capitalization weights $w_{t,n}^{m}$, i.e. $\frac{1}{2}\sum_n |w_{t,n}^{ESG} - w_{t,n}^{m}|$. For all other variables, I report the average across quarters within a given year.

	ESG Fund-Level Statistics					Aggregate Statistics on w_t^{ESG}		
Year	# Funds	Avg. # Stocks	Avg. AUM ($ Billion)	Excess Flows (%)	# Name Changes	AUM ($ Billion)	% Indexed AUM	Active Share
2010	89	95	0.34	0.54	2	30.60	0.11	0.71
2011	82	128	0.38	-1.26	0	31.17	0.13	0.70
2012	88	139	0.28	-0.69	2	25.02	0.16	0.70
2013	83	119	0.34	0.73	1	28.82	0.17	0.69
2014	88	115	0.43	-0.55	3	37.66	0.19	0.67
2015	101	133	0.36	-0.45	3	36.80	0.24	0.68
2016	117	153	0.31	0.35	6	36.62	0.25	0.67
2017	158	153	0.30	2.21	7	48.03	0.21	0.65
2018	199	147	0.32	1.46	19	63.19	0.22	0.63
2019	237	149	0.34	2.09	16	79.71	0.28	0.61
2020	288	165	0.43	3.24	19	126.68	0.40	0.56
2021	368	156	0.63	2.23	21	233.48	0.50	0.57

a sustainable keyword while leaving the fund and portfolio identifier unchanged. The column 'Excess Flows' reports the average flow into ESG funds in excess of the average flow into non-ESG funds. Over the past 5 years, the ESG funds received around 2-3% higher quarterly inflows than other funds. Appendix Section A provides an in-depth analysis of ESG flows controlling for fund characteristics, performance, and portfolio holdings. In a difference-in-difference setting using, I show that having an ESG keyword in the fund-name buys additional quarterly inflows of 1.8%.

B The Representative ESG portfolio

Using Thomson's Mutual Fund Holdings Database (s12 file), I construct the aggregate portfolio held by the sample ESG mutual funds. To this end, let $Q_{t,n}^{ESG} = \sum_{i \in I^{ESG}} Q_{t,n}^{i}$ denote the aggregate holdings of the set of identified ESG funds mutual funds I^{ESG}. The representative ESG mutual fund's portfolio weights are given by ESG funds' total dollar holdings divided by their aggregate assets under management. Formally

$$w_{t,n}^{ESG} = \frac{P_{t,n} Q_{t,n}^{ESG}}{\sum_{n=1}^{N} P_{t,n} Q_{t,n}^{ESG}} \tag{2.1}$$

By using weights instead of dollar holdings, the representative ESG portfolio (henceforth ESG portfolio) is invariant to the number of identified funds, as long as the sample is representative of the average ESG fund. The three rightmost columns of Table 2.1 report summary statistics on the aggregate ESG portfolio. Total assets grew from $30 to $233 billion. At the same time, the fraction of total ESG assets that track an ESG index has also steadily increased to 50%. To what extent do the aggregated holdings across ESG funds reflect the market portfolio? Note, that as more money is flowing into ESG funds, the ESG and the market portfolio converge *by construction*. In the limit, all money is invested in ESG funds and the ESG portfolio coincides with the market portfolio. 'Active Share' is defined as the deviation of the ESG portfolio from the aggregate mutual fund portfolio (henceforth market portfolio). The ESG portfolio tilts around 70% of its assets away from the market portfolio. However, in the most recent years, the active share has declined to 57%.[12] Despite portfolio heterogeneity across ESG funds, their main portfolio tilts go in similar directions. Therefore, while the set of identified ESG funds depends on the kind and amount of keywords used, the aggregate portfolio is extremely robust to different subsets of ESG funds. Appendix Section A provides a detailed investigation of the robustness of the ESG portfolio. Using the ESG portfolio, I construct a revealed-preference measure $\tau_{t,n}$ of investors' green tastes for a stock given by

$$\tau_{t,n} = w_{t,n}^{ESG} - w_{t,n}^{MF} \tag{2.2}$$

where $w_{t,n}^{MF}$ is the aggregate mutual fund portfolio, which is constructed as in (2.1) but summing over all mutual funds instead of the subset of ESG funds. Empirically $w_{t,n}^{MF}$ is extremely close to the market capitalization-weighted portfolio, so that defining $\tau_{t,n}$ in excess of market weights leaves all results of the paper unchanged. Stocks with a higher $\tau_{t,n}$ are perceived to be more sustainable as they are overweighted by the representative ESG portfolio. Note, that the revealed preference measure $\tau_{t,n}$ is also a zero-investment long-short portfolio, that is long $1 in the ESG portfolio and short $1 in the aggregate mutual fund portfolio. I define stocks with $\tau_{t,n} > 0$ as green stocks and stocks with $\tau_{t,n} < 0$ as other (non-green) stocks. This revealed-preference measure is available for all stocks at a monthly frequency over a large time horizon. It furthermore does not rely on subjective sustainability metrics or third-party ESG scores. $\tau_{t,n}$ is therefore a more objective representation of the market's perception of sustainability. Note, that the purpose of this paper is not to identify a measure of *true* sustainability, but to assess the cross-sectional price distortions due to ESG flows. The most adequate measure of sustainability is hence the measure that people implicitly use when they invest sustainably.

In Appendix Section A, I confirm that $\tau_{t,n}$ is robust to the subset of ESG funds used for its

[12] Formally, active share is defined as the deviation of the $\frac{1}{2}\sum_{n \in N^i} |w_{t,n}^{ESG} - w_{t,n}^{MF}|$ where $w_{t,n}^{ESG}$ are the aggregate portfolio weights across all ESG funds and $w_{t,n}^{MF}$ is the aggregate portfolio of all mutual funds.

computation. I compute two different ESG portfolios using random (non-overlapping) subsets of funds and show that the corresponding $\tau_{t,n}$ are highly cross-sectionally correlated ($\rho > 60\%$). A thorough investigation of the difference between *true* and *perceived* sustainability is beyond the scope of this paper, which addresses the distortion of realized returns due to ESG tastes, regardless of whether they are *correct* or not. I nevertheless confirm that $\tau_{t,n}$ is significantly related to commonly used sustainability metrics. The ESG portfolio significantly tilts towards stocks with high ESG scores and underweights sin stocks, stocks in the fossil fuel industry, and high Co2 emitters.[13]

C Realized ESG Returns

Next, I investigate the realized performance of the ESG portfolio $w_{t,n}^{ESG}$, the aggregate mutual fund portfolio $w_{t,n}^{MF}$, and the long-short ESG portfolio $\tau_{t,n} = w_{t,n}^{ESG} - w_{t,n}^{MF}$. The portfolios are rebalanced quarterly based on the funds' SEC filings.[14] Table 2.2 reports the annualized returns and alphas of the portfolios. The first two columns report the annualized returns of

Table 2.2: **ESG Returns**
The table reports annualized average returns and alphas from 2016 to 2021. The two left columns report average annualized returns on the market portfolio w_t^{MF} and the representative ESG portfolio w_t^{ESG}. The three right columns report the annualized alphas of the long-short ESG portfolio $\tau_t = w_t^{ESG} - w_t^{MF}$. The ESG portfolios are rebalanced quarterly. Alphas are computed with respect to the CAPM, the CAPM plus the Green Factor in PST (2022), and the Carhart 4-factor model plus the Green Factor. The standard errors are robust to heteroskedasticity and autocorrelation.

	Mutual Fund Portfolio w_t^{MF}	ESG Portfolio w_t^{ESG}	Long-Short ESG Portfolio ($\tau_{t,n}$) Return	α (CAPM)	α (CAPM + Green)	α (CH4 + Green)
2012-2022						
Return (%)	15.57	16.37	0.72	0.96	0.48	0.42
t-statistic	3.20	3.41	1.57	2.03	1.01	0.90
2016-2022						
Return (%)	16.98	19.11	2.01	2.40	1.87	1.51
t-statistic	2.05	2.36	2.91	3.47	2.55	2.01

the market portfolio and the ESG portfolio. Between 2016 and 2021 the ESG portfolio had a significant 2% higher annualized return than the market portfolio. The four right columns report the returns and alphas of the long-short ESG portfolio $\tau_t \in \mathbb{R}^N$. The returns of τ_t will henceforth be referred to as ESG returns. Intuitively, one would expect significantly negative alphas capturing the taste premium investors are willing to give up in order to invest according

[13] See Appendix Section A for details.
[14] The portfolios are not necessarily tradeable as funds usually delay their SEC report by up to 45 days.

to their ESG preferences. However, long-short returns and alphas are significantly positive. From 2016 to 2021, the long-short ESG portfolio had a significant annual CAPM alpha of 2.4%.[15] The last column controls for PST's (2022) Green Factor, as well as the Carhart 4-Factors. The alpha merely drops to 1.5% and remains significant with t-Statistic of 2.01. Overall, Table 2.2 suggests that investors have been *rewarded* instead of *penalized* for investing according to their ESG preferences. The weights in the long-short ESG portfolio τ_t can be interpreted as a measure of investors' perception of sustainability. Thus, regardless of their *true* sustainability, the stocks that investors deemed more sustainable tended to have higher returns than others.

Despite the apparent outperformance of the ESG portfolio, the goal of this paper is not to add to the debate about whether or why sustainable investing has higher or lower *expected* returns in equilibrium. This paper tries to answer the question of how, ceteris paribus, the cross-section of *realized* returns responds to flows to the ESG portfolio. Thus we can assess, to what extent the realized returns from sustainable investing have been driven by flows towards sustainable funds. However, total flows in the ESG portfolio are not directly observable. The next section shows how to construct aggregate ESG flows from institutional portfolio holdings.

D Measuring Total Flows in the ESG Portfolio

Total flows into the ESG portfolio are difficult to observe. According to Morningstar, labelled ESG mutual funds held \$350 billion in total assets, which was less than 1% of the total \$37 trillion held by all ETFs and Mutual funds in the US.[16] However, this does not include the (unobservable) ESG tilts of other mutual funds, investment advisors, pension funds, banks, insurance companies, and other institutions. Therefore, the flows into labelled ESG mutual funds only represent a small subset of total ESG flows. In order to get a sense of total ESG flows, I use 13F filings to estimate each institution's 'ESG share'. Here, I merely present the main procedure. Technical details are delegated to Appendix B. For simplicity, I omit the institution and quarter labels i and t. I project each 13F institution's portfolio w_n in the cross-section onto a set of S managed portfolios w_n^s,

$$w_n = \sum_{s=1}^{S} \beta^s w_n^s + a_n \tag{2.3}$$

where $n \in N_t^i$ is the subset of stocks held by i as of quarter t. The managed portfolios are constructed such that the weights add up to 1 across all stocks currently held by the institution. For example, the 'managed' market-weighted portfolio ($s = Mkt$) is given by $w_n^{Mkt} = P_n / \sum_{n \in N^i} P_n$. The residual a_n from the projection is an active zero-cost long-short portfolio in the spirit of

[15] Note, that these are not the *true* returns an investor would have achieved by investing in the asset-weighted portfolio of (ESG) mutual funds because of fees and because many of these funds trade actively within quarters.

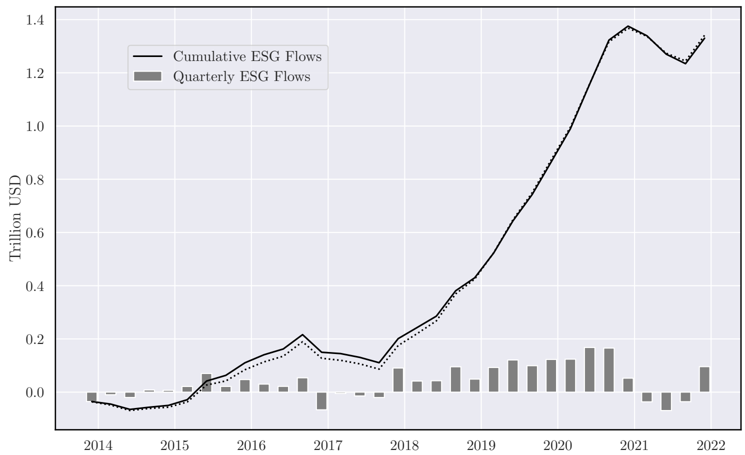
[16] See Morningstar's 2021 Sustainable Funds U.S. Landscape Report. The assets of labelled ESG funds from the previous section are of a similar magnitude.

Cremers and Petajisto (2009).[17] Thus the coefficients β_t^s sum to 1 and can be interpreted as the asset shares of managed portfolios within institution i. The coefficient on the ESG-managed portfolio, β^{ESG}, measures the institution's ESG share. I then compute the total ESG flow as the sum of institution-specific ESG flows

$$F_{t+1}^{ESG} = \sum_{i=1}^{I} A_{t+1}^{i,ESG} - A_t^{i,ESG}(1 + R_{t+1}^{ESG}) \tag{2.4}$$

where $A_t^{i,ESG} = \beta_t^{i,ESG} A_t^i$ are the total assets of institution i allocated to the ESG portfolio at time t and R_{t+1}^{ESG} is the return on the ESG portfolio. This measure of total ESG flows is highly robust to controlling for different managed portfolios in the estimation of $\beta_t^{i,ESG}$.[18] Figure 2.1 plots the total flow into the ESG portfolio from 2012 to 2022. Total ESG flows have increased

Figure 2.1: **Total ESG Flow**
The figure plots the total flow into the ESG portfolio from 2012 to 2022. I compute the ESG flow for each 13F institution as the return-adjusted change in ESG-assets under management and then then sum across all institutions. I report rolling 4-quarter averages and plot the cumulative sum of all flows since 2012. The dotted line plots the ESG flow when controlling for exposures to 12 (Fama-French) industry portfolios in the estimation of β_t^{ESG}.



rapidly since 2017 and amount to approximately $1.3 trillion as of 2022, which far exceeds the flows into explicitly labelled ESG mutual funds. Having constructed total ESG flows, we are now in the position to assess their impact on the realized returns of the ESG portfolio. The key difficulty in measuring flow-driven price impact lies in the joint endogeneity of prices and demand, which prevents simple regressions of realized returns onto flows. The next section introduces a structural approach to estimating the price impact of ESG flows.

[17] In fact, $1/2 \sum_{n \in N^i} |a_n|$ and the active share in Cremers and Petajisto (2009) coincide if $a_n = w_n - w_n^{Mkt}$. This is the case if the coefficient on the market portfolio β^{Mkt} is equal to 1 and the coefficients on all other managed portfolios are equal to 0.

[18] See Appendix B for details.

2.3 A Structural Model of Price Pressure

A Setup and Variable Definitions

This section provides a structural approach to estimating the link between demand shocks and prices. The setup closely follows van der Beck (2022). Here, I merely state the variables and main structural estimation equations. There are N stocks indexed by $n = 1,...,N$ and T time periods $t = 1,...,T$. Shares outstanding are normalized to 1 such that the price of a stock, $P_{t,n}$, coincides with market equity. Lowercase letters denote logs (if not otherwise specified) and one-period changes in variables are denoted by $\Delta x_t = x_t - x_{t-1}$. There are I investors indexed by $i = 1,...,I$ that hold a subset $N_t^i \subseteq N$ of all stocks. $Q_t^i \in \mathbb{R}^{N^i}$ denotes the vector of shares held by i. Because of the normalization, Q_t^i are equal to ownership shares such that $\sum_{i=1}^I Q_{t,n}^i = 1$. The optimal portfolio $Q_t^i = f^i(P_t, V_t)$ is a function of the vector of current stock prices $P_t \in \mathbb{R}^N$ and a collection of other exogenous observable and unobservable variables V_t (such as the assets under management, interest rate, fundamentals, or investment constraints). An investor's elasticity of demand with respect to the price (henceforth elasticity of demand) is defined as the negative percentage change in holdings when the price of a stock increases by 1 %. Formally,

$$\zeta_{t,n}^i = -\frac{\partial Q_{t,n}^i / Q_{t,n}^i}{\partial P_{t,n} / P_{t,n}} \tag{2.5}$$

Similarly, the cross-elasticity of demand is given by $\zeta_{t,nm}^i = -\frac{\partial Q_{t,n}^i / Q_{t,n}^i}{\partial P_{t,m} / P_{t,m}}$ and measures how much of n investor i sells when m's price increases by 1%.[19] The stock-specific and cross elasticities can be stacked in an $N^i \times N^i$ elasticity matrix ζ_t^i for every investor. The aggregate elasticity of demand is defined as the ownership-weighted sum of the investor-specific elasticity matrices,

$$\zeta_t = \sum_{i=1}^I \text{diag}(Q_t^i)\zeta_t^i \tag{2.6}$$

with elements equal to $\zeta_{t,nm} = \sum_{i=1}^I Q_{t,n}^i \zeta_{t,nm}^i$.[20] Stocks that are primarily held by passive index funds (with $\zeta_t^i = 0$) have a low aggregate elasticity. The distribution of ownership, therefore, affects the aggregate elasticity. For example, the rise of passive investing increases the ownership of less elastic investors which drives down the aggregate elasticity, unless the active investors substantially increase their elasticity (see Haddad et al. (2021)).

Lastly, let $\Delta d_{t,n}$ denote a demand shock for n between t and $t+1$, expressed as a fraction of shares outstanding. The demand shock could be flow-induced purchases of green stocks by an ESG fund or the inclusion of a stock in an ESG index and the corresponding purchases by index trackers.

[19] Note, that for $m = n$ the cross-elasticities $\zeta_{t,nn}^i$ are equal to stock-specific elasticities $\zeta_{t,n}^i$.

[20] Recall, that shares outstanding are normalized to 1. Therefore $\sum_{i=1}^I \text{diag}(Q_t^i)$ is equal to the identity matrix.

B Demand-Driven Price Impact

Now assume that an ESG fund receives large inflows and proportionally expands its existing positions resulting in an exogenous demand shock $\Delta d_t \in \mathbb{R}^N$. Equilibrium prices adjust in order to accommodate the demand shock resulting in realized log returns $\Delta p_t \in \mathbb{R}^N$. Proposition 1 in van der Beck (2022) shows that a first order approximation to Δp_t for a large class of models (including e.g. the CAPM or Demand System Approach to Asset Pricing) is given by

$$\Delta p_t = \mathcal{M}_t \Delta d_t + \epsilon_t \tag{2.7}$$

where $\mathcal{M}_t \in \mathbb{R}^{N\times N}$ is a price pressure matrix equal to the inverse of the market's aggregate elasticity of demand

$$\mathcal{M}_t = \zeta_t^{-1}. \tag{2.8}$$

See van der Beck (2022) for a proof. ϵ_t captures other sources of return variation such as factor exposures or fundamental news and is orthogonal to the demand shock Δd_t. As the focus of this paper is purely empirical, equation (2.7) can also be viewed as an assumption as in Greenwood and Thesmar (2011). The link between demand shocks and prices is given by the inverse of the market's elasticity of demand $\mathcal{M}_t$, henceforth referred to as the multiplier matrix. The more elastic investors are (i.e. the larger the diagonal elements in ζ_t^i), the less prices of green stocks have to move, in order to accommodate the demand shocks from flows to ESG funds. Cross-elasticities drive the off-diagonal elements in $\mathcal{M}_t$ and are responsible for flow-induced spill-over effects to other stocks. If investors accommodate flow-driven price pressure on green stocks primarily by substituting towards brown industries, the relative price impact of ESG investing may be negligible.

Example. In order to bolster intuition for the importance of cross-elasticities, consider the following simplified example: There are two stocks, a green stock g and a brown stock b with a market capitalization of 1, and a representative investor with a 2×2 elasticity matrix. Her demand elasticities with respect to g and b are the same, i.e. $\zeta_g = \zeta_b$. Also, her elasticity of substitution is the same moving from g to b and vice versa, i.e. $\zeta_{b,g} = \zeta_{g,b}$. Now assume that there is an exogenous ESG flow in g and b equal to \$1 and -\$1 respectively. The flow-driven price pressure (2.7), is given by $\begin{bmatrix} \zeta_g & -\zeta_{g,b} \\ -\zeta_{g,b} & \zeta_g \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. The difference in market capitalization of g and b after the demand shock is given by $\frac{2}{\zeta_g + \zeta_{g,b}}$.[21] First, the greater the stock-specific elasticity (i.e. the more willing the investor is to sell green and buy brown shares) the smaller the price impact. Second, the greater the cross-elasticity (i.e investors' substitution towards brown stocks as a result of the price increase of the green stock) the smaller the equilibrium price impact. The equilibrium impact of ESG investing, therefore, depends on i) how willing the arbitrageurs are to provide green shares and ii) which stocks they substitute towards.

[21] To see this, note that the log return on green and brown stocks is given by $\begin{bmatrix} \zeta_g & -\zeta_{g,b} \\ -\zeta_{g,b} & \zeta_g \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Multiplying by market equity (which is equal to 1) approximates the change in dollar terms. The difference in market equity between the green and the brown stock after the flow is therefore given by $[1 \; -1] \begin{bmatrix} \zeta_g & -\zeta_{g,b} \\ -\zeta_{g,b} & \zeta_g \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{2(\zeta_g - \zeta_{g,b})}{\zeta_g^2 - \zeta_{g,b}^2} = \frac{2}{\zeta_g + \zeta_{g,b}}$.

C Structural versus Reduced-Form Estimation

As outlined above (and expressed in detail in van der Beck (2022)) the matrix $\mathcal{M}_t$, which links demand shocks and the cross-section of realized returns, can be obtained structurally from investors' demand elasticities. Before diving into estimating elasticities from holdings data, it is worth stepping back and asking whether a structural estimation is truly necessary. One could imagine a much simpler identification from directly regressing realized ESG returns onto demand shocks similar to PST (2022). For example, Pavlova and Sikorskaya (2022) regress returns onto changes in benchmarking intensities and obtain a multiplier of around 1.5. After all, estimating demand elasticities via regressions of demand onto prices is subject to the same endogeneity concerns that contaminate regressions of prices onto demand: Both are jointly determined in equilibrium. Assume, that we had access to non-fundamental demand shocks for green stocks Δd_t from e.g. a stock's inclusion in an ESG index as in Berk and van Binsbergen (2022).[22] The shocks could be used to directly estimate the multiplier using (2.7) as a linear regression. Nevertheless, there are three distinct benefits of the structural approach. First, it gives rich insights into the underlying investor-specific determinants of the flow multiplier. Second, one can obtain stock-specific and time-varying effects even if we estimate a scalar elasticity for every investor. Because the ownership $Q^i_{t,n}$ varies across stocks and time, ownership-weighted sums across elasticities $\sum_{i=1}^{I} Q^i_{t,n}\zeta^i$ vary across stocks and time. Third, one can use a large cross-section of holdings data over a long history to identify ζ_t as opposed to the small number of potential ESG demand shocks.

Lastly, note that the elasticities themselves are not *deep* parameters and could be a function of trading costs, risk aversion, or investment constraints. The model and its estimation are therefore 'semi-structural'. Understanding the drivers of demand elasticities and in particular downward-sloping demand curves is an important avenue for future research.

D Estimating Elasticities

As in van der Beck (2022), I define institutions' demand as $\Delta q^i_{t,n} = \log Q^i_{t,n} - \log Q^i_{t-1,n} \approx \Delta Q^i_{t,n}/Q^i_{t-1,n}$. Thus $\Delta q^i_{t,n}$ simply measures the percentage change in shares held by institution i in stock n between two quarters. Similarly, percentage changes in the price are given by $\Delta p_{t,n} = \log P_{t,n} - \log P_{t-1,n} \approx \Delta P_{t,n}/P_{t-1,n}$. These variable definitions directly emerge from the definition of the elasticity (2.5). Up to a first order, an investor's demand elasticity can be written as a linear regression of trades $\Delta q^i_{t,n}$ onto log returns $\Delta p_{t,n}$.

$$\Delta q^i_{t,n} = -\zeta^i \Delta p_{t,n} + \epsilon^i_{t,n} \tag{2.9}$$

where $\epsilon^i_{t,n}$ captures demand shocks due to e.g. fundamentals, flows or trading constraints. The reduced-form specification essentially corresponds to a first difference estimator of the

[22]See Shleifer (1986), Coval and Stafford (2007) or Schmickler and Tremacoldi-Rossi (2022) for other examples of non-fundamental demand shocks

logit demand specification in KY (2019). van der Beck (2022) provides a detailed investigation of the relationship between the two estimators, which is summarized in Appendix D of this paper. Note, that the scalar regression coefficient ζ^i is a reduced-form approximation of an investor's elasticity, which does not ensure that the investor's total assets remain unchanged. Appendix D shows how to incorporate (2.9) in a logit framework that satisfies the budget constraint and allows constructing the full time-varying elasticity matrix $\zeta_t^i \in \mathbb{R}^{N\times N}$ from the scalar regression coefficient ζ^i and portfolio holdings.

E Identification

A causal identification of demand elasticities requires exogenous variation in prices that is orthogonal to the investor's own demand shocks. In other words, we can use the exogenous demand shocks of one investor to identify the elasticity of another investor. As for every buyer there is a seller, exogenous demand shocks by one investor can essentially be viewed as shifting the supply curve. The literature has proposed a variety of potential instruments such as index inclusions, mutual fund flows or dividend reinvestments.[23] An advantage of estimating elasticities via trades (instead of holdings as in KY (2019)) is that essentially all of the instruments from the event-study literature on price pressure can be re-employed to identify demand elasticities. Van der Beck (2022) uses flow-driven trades by mutual funds as an exogenous shock to identify elasticities. Many mutual funds scale their existing holdings in response to in- and outflows (see Lou (2012)). Aggregating the flow-driven trades across all mutual funds provides exogenous cross-sectional demand shocks under the (strong) assumption that the flows were not driven by the funds' underlying fundamentals. To address these concerns, van der Beck (2022) constructs surprise flows by orthogonalizing the cross-section of mutual funds flows with respect to the funds' underlying holdings and characteristics.[24] However, it remains unclear whether a simple orthogonalization provides true exogenous flow shocks. In this paper, I take one step further and construct exogenous flow shocks from dividend reinvestments as in Schmickler and Tremacoldi-Rossi (2022). I closely follow their construction of dividend-induced mutual fund trades. Let $D_{t,n}$ denote stock n's dividends per share paid in quarter t. For every fund i, I construct dividend flow df_t^i as the total dividend payout across all stocks in the portfolio relative to assets under management:

$$df_t^i = \sum_{n\in N^i} D_{t,n} Q_{t-1,n}^i / A_{t-1}^i \tag{2.10}$$

In appendix 2.13, I show that mutual funds tend to proportionally reinvest aggregate dividend payouts in their existing portfolios.[25]. The hypothetical trading in stock n due to reinvested dividend flows is given by $df_t^i Q_{t-1,n}^i$. I construct an instrument for each investor i by summing

[23] See Chang et al. (2015), Lou (2012), and Schmickler (2020) for respective examples.

[24] Similarly, Schmickler (2020) constructs high-frequency flow shocks to address contemporaneous return chasing in mutual fund flows.

[25] Chen (2020) arrives at a similar result

the dividend-induced trades (DIT) by all other mutual funds,

$$DIT_{t,n}^{-i} = \sum_{j \in MF, j \neq i} df_t^j Q_{t-1,n}^j \tag{2.11}$$

Note, that the dividend announcement date of stock n, which contains fundamental information, often lies in the same quarter as the dividend payment. To avoid including the fundamental news coming from n's own dividend announcement, I construct $DIT_{t,n}$ using $df_{t,-n}^i = \sum_{m \neq n} D_{t,m} Q_{t,m}^i / A_t^i$ instead of df_t^i.[26]

Having constructed investor-specific instruments, the elasticities can be obtained in a simple two-stage least squares procedure. Let $\Delta\hat{p}_t^i$ denote the fitted value from regressing returns onto the investor-specific instrument $f_{t,n}^{-i}$. The second stage regression of investor-specific trades Δq_t^i onto the investor-specific instrumented return $\Delta\hat{p}_t^i$ allows identifying their demand elasticities ζ_t^i. Formally, for every investor the two stages are given by:

$$\begin{aligned} \text{1st Stage:} \quad & \Delta p_{t,n}^i = \theta^i DIT_{t,n}^{-i} + \epsilon_{t,n}^i \\ \text{2nd Stage:} \quad & \Delta q_{t,n}^i = -\zeta^i \Delta\hat{p}_{t,n}^i + \epsilon_{t,n}^i \end{aligned} \tag{2.12}$$

where $\epsilon_{t,n}^i = \sum_{k=1}^K X_{t,n,k}\beta_k^i + u_{t,n}^i$ includes the control variables log book equity, profitability, investment, and market beta. The trading due to aggregate dividend flows $DIT_{t,n}^{-i}$ is plausibly more exogenous than ordinary flow-induced trading. The drawback of this instrument is, however, that we cannot obtain negative demand shocks as dividends are strictly positive. Thus the identified elasticities only capture how stock price *increases* affect demand. As a robustness check, Appendix C reports the estimated elasticities identified from flow shock-induced trading, which can take on both positive and negative values. I also explore the stability of the estimates by using changes in 'Benchmarking Intensity' (BMI) by Pavlova and Sikorskaya (2022) as an alternative instrument in the first stage. The next section and Appendix Section C provide further details.

F The Multiplier Matrix

I estimate ζ^i over the panel of quarterly holdings from 2010 to 2021.[27] The multiplier matrix $\mathcal{M} \in \mathbb{R}^{N\times N}$ is given by the inverse of the aggregate (ownership-weighted) elasticity. I omit the time t subscript for notational simplicity. The diagonal elements of $\mathcal{M}$ are the stock-specific multiplier effects. The n-th diagonal element $\mathcal{M}_{n,n} = \frac{\Delta p_n}{\Delta d_n}$ measures the price impact of demand shocks for n onto the price of n. The off-diagonal elements are the spillover effects to other stocks. In particular, $\mathcal{M}_{m,n} = \frac{\Delta p_m}{\Delta d_n}$ measures the price impact of demand shocks for n onto the price of m. Let $N^G \subset N$ denote the subset of green stocks. We are interested in the price impact of demand shocks for green stocks N^G onto the cross-section of all stocks N. Omitting the time subscript, one can partition the multiplier matrix into submatrices by green

[26] See Schmickler and Tremacoldi-Rossi (2022).

[27] The estimated investor-specific coefficients are reported in Appendix Table 2.16.

($g \in N^G$) and other ($b \notin N^G$) stocks as

$$\mathcal{M} = \begin{bmatrix} \mathcal{M}_{gg} & \mathcal{M}_{gb} \\ \mathcal{M}_{bg} & \mathcal{M}_{gb} \end{bmatrix}. \tag{2.13}$$

The important elements are $\mathcal{M}_{gg}$ and $\mathcal{M}_{bg}$, which capture the effect of green demand shocks onto green stocks (gg) and of green demand shocks onto non-green stocks (bg) respectively. [28] Intuitively, $\mathcal{M}_{gg}$ measures by how much the prices of other green stocks go up when the demand for any green stock increases by 1%. $\mathcal{M}_{bg}$ measures by how much the prices of non-green stocks increase. The diagonal elements of $\mathcal{M}_{gg}$ are the direct stock-specific effects of green demand, i.e. the price increase of n as a response to a 1% demand shock for n.[29] The cross-multipliers among green stocks are a key determinant of the spillover effects of ESG demand. If market participants accommodate the demand for green stocks by substituting towards other green stocks, the relative repricing of green versus brown stocks due to ESG flow may be strongly amplified.

Table 2.3 summarizes the elements of the multiplier matrix $\mathcal{M}_t$. The first column reports the direct impact of ESG demand, i.e. the diagonal elements of $\mathcal{M}_{gg}$. The remaining columns report the cross-multipliers, i.e. the spillover effects onto other green and non-green stocks.

Table 2.3: **The Elements of Multiplier Matrix**
The table summarizes the stock-specific and cross-elements of the multiplier matrix for green demand shocks. The first column reports the stock-specific multipliers, i.e. the percent increase in price of a green stock following a 1% increase in demand for that stock. The other columns report the off-diagonal elements of the multiplier matrix, i.e. cross-multipliers, which capture spillover effects to other stocks. Cross-multipliers are separated into spillover effects within green stocks ($\mathcal{M}_{gg}$) and from green stocks to brown stocks ($\mathcal{M}_{bg}$).

		Cross-Multipliers ($\times 10^4$)	
	$\mathcal{M}_g$	$\mathcal{M}_{gg}$	$\mathcal{M}_{bg}$
Mean	1.11	-0.86	-1.40
Std.	(0.1)	(9.66)	(8.40)
10th Pctl.	1.01	-2.30	-3.23
Median	1.09	-0.05	-0.12
90th Pctl.	1.25	0.43	0.36
Fraction Positive Spillovers		38%	32%

The average multiplier of the demand for green stocks is around 1.11, implying that (on

[28]Formally, letting $N^B = N - N^G$ denote the number of non-green stocks, the dimensions of the matrices are given by $\mathcal{M}_{gg} \in \mathbb{R}^{N^G \times N^G}$, $\mathcal{M}_{bb} \in \mathbb{R}^{N^B \times N^B}$, $\mathcal{M}_{gb} \in \mathbb{R}^{N^G \times N^B}$ and $\mathcal{M}_{bg} \in \mathbb{R}^{N^B \times N^G}$.
[29]Note, that because ownership shares $Q^i_{t,n}$ vary across stocks, the elasticity matrix is not symmetrical. Therefore $\mathcal{M}_{gb}$ and $\mathcal{M}_{bg}$ are different objects.

average) a 1% increase in the demand for a green stock leads to a 1.11% increase in the price of that stock. The stock-specific multiplier is positive for all stocks because demand is downward-sloping for all investors.[30] This is the key channel through which continued capital flows into green firms can lead to high realized returns. Among the cross-multipliers, there is great heterogeneity across stocks which implies that the spillover effects of ESG demand shocks are highly nontrivial. Notably, there are more positive spillover effects of ESG demand towards other green stocks than towards non-green stocks. On average, a positive demand shock for a green stock leads to a price increase for roughly 40% of all other green stocks.

How sensitive are the multiplier estimates to an alternative identification? In Appendix Section C I identify investor-specific elasticities using an alternative instrument, namely changes in benchmarking intensity (BMI) by Pavlova and Sikorskaya (2022). A stock's BMI measures the fraction of the total market capitalization held by benchmarked investors. Changes in BMI reduce the effective supply of a stock and can be used to identify investor-specific elasticities. The multipliers obtained from the BMI-based elasticities are of strikingly similar magnitude. Using the alternative identification, a 1% demand shock for the average green stock raises its price by 1.17%. As a final robustness check, Appendix Section C estimates investor-specific elasticities from surprise-flow induced trading as in van der Beck (2022). Using surprise-flows as an instrument, I estimate that a 1% demand shock for the average green stock raises its price by 2.78%.

2.4 The Aggregate Impact of ESG Flows

Having estimated the market's willingness to accommodate ESG demand we are now in the position to estimate the impact of flows on the realized returns from ESG investing.

A ESG Flow Multiplier

What is the impact on valuations, if investors reallocate \$1 from the market portfolio towards the ESG portfolio? A \$1 ESG flow translates into stock-specific demand shocks given by $\tau_{t,n} = w^{ESG}_{t,n} - w^{MF}_{t,n}$. Equation (2.7) then implies, that the equilibrium change in prices $\Delta P^{ESG}_{t+1} \in \mathbb{R}^N$ due to ESG flows is simply

$$\Delta P^{ESG}_{t+1} = \mathcal{M}_t \tau_t. \tag{2.14}$$

Note, that (2.7) is expressed in percentage terms, i.e. the return $\Delta p_{t+1,n}$ resulting from a demand shock in percent of shares outstanding. It can also be expressed in terms of dollar terms by multiplying by prices $P_{t,n}$ (which are equal to market equities due to the normalization). Here, net flows are equal to zero as $\sum_{n=1}^{N} \tau_{t,n} = 0$. One could alternatively model nonzero net

[30] Unlike in KY (2019), this is not an assumption. The estimation in changes yields downward-sloping demand curves for *all* investors without a coefficient constraint. See van der Beck (2022) for details.

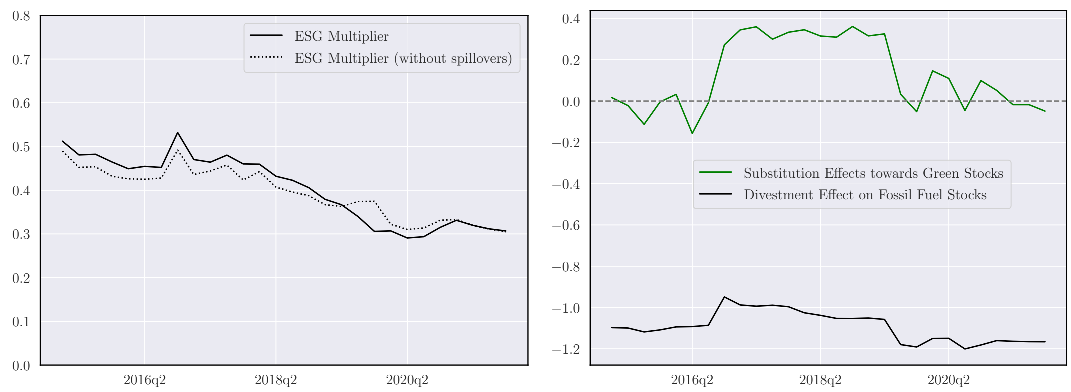
equity flows, as inflows to ESG funds could also come from e.g. households that were not previously invested in the stock market. Such flows would affect both aggregate stock market and ESG returns. As the focus of this paper lies on the *excess* returns of ESG funds (over the aggregate mutual fund portfolio), net-zero flows are a more suitable way of modelling ESG demand.

Summing the flow-induced change in market equity across all green stocks yields the aggregate dollar impact of a one-dollar ESG flow on all green firms (or at least the ones perceived to be green).[31] This effect will henceforth be referred to as the ESG flow multiplier. The ESG flow multiplier is driven by two components: First, flows play a stronger role in cross-sectionally inelastic markets with a low aggregate demand elasticity for green stocks and therefore a high multiplier matrix $\mathcal{M}_t$. Intuitively, if price inelastic investors (i.e. investors with a low $\zeta_{t,n}$) are the main shareholders of green stocks (i.e. they have a high ownership $Q^i_{t,n}$), then aggregate elasticity for green stocks is low and prices have to adjust a lot in order to accommodate flow-induced demand. Second, the impact of ESG flows depends on the deviation of the ESG portfolio from the market portfolio $\tau_{t,n}$. If ESG funds' deviation from the aggregate mutual fund portfolio is negligible, then flows towards sustainable funds have no impact on the price regardless of the multiplier effect $\mathcal{M}_t$. Panel (a) of Figure 2.2 plots the ESG flow multiplier

Figure 2.2: **ESG Flow Multiplier**
The figure plots the ESG flow multiplier, i.e. the aggregate change in market cap in green stocks due to a $1 ESG flow. Formally, the aggregate effect of ESG flows on green stocks is given by $\sum_{n \in N^G} \Delta P^{ESG}_{t+1,n}$ where $\Delta P^{ESG}_t = \mathcal{M}_t \tau_t$ is the vector of price changes following the $1 ESG flow. The dotted line reports the ESG multiplier without cross-spillover effects (setting the diagonal elements in $\mathcal{M}$ to 0. Panel (b) plots the impact of a divestment strategy that divests $1 from a value-weighted portfolio of all fossil companies. The black line shows the direct impact on the aggregate valuation of fossil fuel companies. The green line shows the indirect spillover effects to green stocks.

(a) ESG Flow Multiplier

(b) Fossil-Fuel Divestment Multiplier



over time. The ESG flow multiplier is around 0.5 and has declined to 0.3 in recent years. Thus, withdrawing $1 from the market portfolio and investing it in the ESG portfolio leads to an increase in green stocks' aggregate market capitalization of around $0.3-0.5. The decline in

[31] Let $N^G_t \subset N$ denote the subset of green stocks (for which $\tau_{t,n} > 0$). The total impact on green stocks is then given by $\sum_{n \in N^G_t} \Delta P^{ESG}_{t+1,n}$ where $\Delta P^{ESG}_{t+1,n}$ are the stock-specific entries of ΔP^{ESG}_{t+1}.

the ESG multiplier is directly related to the decline in the active share of the ESG portfolio in recent years. As the ESG portfolio moves closer to the market portfolio, net zero ESG flows lead to smaller demand shocks and therefore a smaller price impact. While arguably more objective than third-party ESG scores, using $\tau_{t,n}$ as a measure of sustainability remains a subjective choice. In order to provide a broader perspective on the efficacy of ESG investing, I also compute the impact of a divestment strategy that divests \$1 from a value-weighted portfolio of all fossil companies. Panel (b) of Figure 2.2 plots the impact of the divestment strategy on the aggregate market capitalization of fossil fuel and green companies. Every dollar withdrawn from the fossil fuel industry reduces its aggregate market capitalization by \$1-1.2. These estimates suggest that divestment strategies can have a large effect on stock prices and therefore firms' cost of capital. Even though green stocks are not directly affected by the divestment strategy, their aggregate value is affected via spillover effects. This underlines the importance of accounting for the off-diagonal elements in $\mathscr{M}_t$. As market participants accommodate the demand shock from the divestment strategy by buying fossil fuel companies they simultaneously buy green companies (potentially to maintain a constant industry exposure). This exerts price pressure on the latter resulting in positive spillover effects.

B Counterfactual ESG Returns in the Absence of Flows

The ESG flow multiplier paired with the large ESG flows of \$1.3 trillion suggest that the flow-driven demand for green stocks has potentially large aggregate pricing implications. In order to assess the quantitative return distortion from total ESG flows, F_{t+1}^{ESG}, I conduct a simple simulation. I simulate counterfactual realized ESG returns if the quarterly flows to ESG funds F_{t+1}^{ESG} were instead reinvested in the aggregate mutual fund portfolio w_t^{MF}.[32] Table 2.4 reports the counterfactual ESG returns in the absence of flow-driven price pressure. The first row reports the empirically observed annualized ESG return, which is defined as the excess return of the ESG portfolio over the aggregate mutual fund portfolio $\tau_t = w_t^{ESG} - w_t^{MF}$. The second row reports the counterfactual ESG return without price pressure from flows towards labelled ESG mutual funds. The raw return and alphas drop by merely 10 basis points. The impact of capital flows towards specifically labelled sustainable mutual funds is therefore negligible. Without the price pressure from *total* ESG flows, however, the raw return and alphas drop by 200 basis points and are all zero. Thus, when assessing the impact of ESG investing, it is important to account for the ESG tilts by all institutions, including large investment advisors, banks, and pensions funds. The results emphasize the sizeable gap between realized and expected returns from ESG investing that is driven by total sustainable flows. Taking the estimates at face value, this suggests that without a continued flow to the ESG portfolio, ESG investing does not have positive abnormal returns. In other words, it is the price pressure from

[32]A first order approximation of the simulated price pressure from $\$X$ net-zero ESG flows is given by $\Delta P_t^{\text{sim}} = \mathscr{M}_t(w_t^{ESG} - w_t^{MF}) * \$X$. The counterfactual returns in the absence of price pressure are then given by $r_{t,n}^{cf} = r_{t,n} - \Delta P_{t+1,n}^{\text{sim}} / P_{t,n}$. Counterfactual ESG returns are then given by $\sum_{n=1}^{N} r_{t,n}^{cf} \tau_{t,n}$. See Appendix D for details.

Table 2.4: **Counterfactual ESG Returns without Flow-Driven Price Pressure**
The table reports the true (empirically observed) realized returns of the long-short ESG portfolio τ_t and the counterfactual returns observed in the absence of price pressure from i) labelled ESG mutual fund flows and ii) total ESG flows. I report raw returns and alphas with respect to the CAPM, the CAPM plus the Green Factor from PST (2022), and the Carhart 4-Factor Model plus the Green Factor.

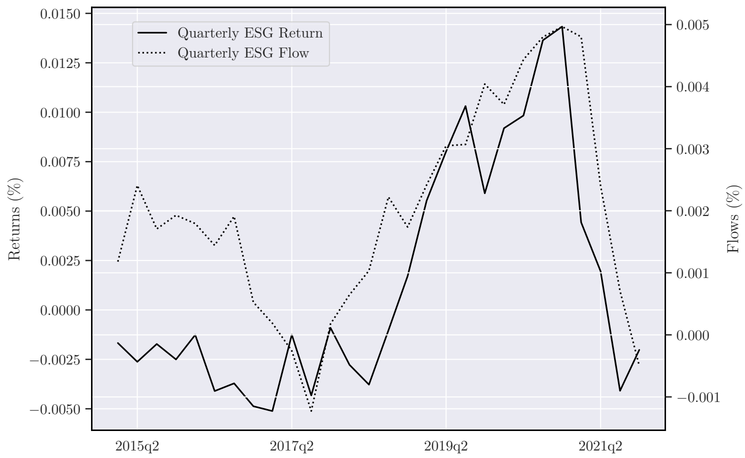
	Return	α (CAPM)	α (CAPM + Green)	α (CH4 + Green)
True Returns: Empirically Observed				
Return (%)	2.01	2.40	1.87	1.51
t-statistic	2.91	3.47	2.55	2.01
Counterfactual Returns: In Absence of Flows from labelled ESG Mutual Funds				
Return (%)	1.92	2.32	1.78	1.42
t-statistic	2.78	3.35	2.43	1.90
Counterfactual Returns: In Absence of Total ESG Flows				
Return (%)	0.04	0.57	-0.05	-0.30
t-statistic	0.05	0.77	-0.07	-0.38

ESG flows that made 'doing well by doing good'-investing possible.

C ESG Flows and Returns: Reduced-Form Evidence

The structural approach presented above allows for circumventing the issue that flows and returns are jointly endogenous. Within the model, ESG flows have a large impact on ESG returns. This result is based on three findings: Large flows towards the ESG portfolio (F_{t+1}^{ESG}), a low elasticity of substitution between green and brown firms ($\mathscr{M}_t$), and a considerable deviation of the ESG portfolio from the market portfolio (τ_t). If ESG returns are truly flow-driven, then aggregate ESG flows should be correlated to realized ESG returns. Figure 2.3 plots the quarterly flow into the ESG portfolio along with the excess return on the ESG portfolio. The correlation between quarterly ESG flows and returns is 74%. While this correlation is by no means causal evidence for flow-driven price pressure, it is nevertheless strikingly high. Notably, PST (2022) find that ESG returns and flows into labelled ESG mutual funds are not significantly correlated. Table 2.5 replicates their result and provides further evidence on the potential importance of flows in explaining ESG returns. I first regress ESG returns on the total ESG flow and the flow to labelled ESG mutual funds separately. Both measures of ESG flows are significantly related to ESG returns with an R^2 of 29%. Note that simple regressions of the returns onto flows cannot identify price pressure: Beliefs about the climate, the fundamentals of ESG firms, and positive feedback trading, drive both flows into ESG funds, as well as the

Figure 2.3: **Aggregate ESG Flows and Returns**
The figure plots the quarterly excess return on the ESG portfolio $\tau_t = w_t^{ESG} - w_t^{MF}$ against the quarterly ESG flow F_{t+1}^{ESG} measured in percent relative to the total stock market capitalization. I plot rolling 4-quarter averages of returns and flows.



return on their underlying assets. I merely present these correlations as suggestive evidence for a potential link between ESG flows and returns. The second set of columns replicates the findings of PST (2022). I regress their GMB factor return onto total ESG flows, ESG mutual fund flows, and instrumented flows using quarterly lags. Confirming their results, I find no significant relationship between GMB returns and ESG flows. This underlines the importance of computing the suitable flow into the object of interest. It is unclear whether flows to ESG mutual funds are indeed directed at the GMB portfolio. While many investors follow the MSCI ESG ratings used in PST (2022), the direction of ESG flows critically depend on how the ratings are used to construct portfolio weights. Thus ESG flows may not directly target the GMB portfolio. I circumvent this issue by investigating flows and returns of the same portfolio (w_t^{ESG}).

2.5 The Cross-Sectional Impact of ESG Flows

This section puts the hypothesis of flow-driven ESG returns to a stronger test. If flow-driven purchases by institutions drive aggregate ESG returns, they should also affect the cross-section of ESG returns. In other words, green stocks that experience higher flow-driven demand should exhibit higher realized returns in the cross-section.

Table 2.5: **The Correlation of ESG Flows and Returns**
The table reports regressions of the form

$$R_t^{ESG} = \alpha + \beta F_t^{ESG} + \epsilon_t$$

where R_t^{ESG} is an ESG return and F_t^{ESG} a measure of ESG flows. The first set of columns uses the τ-portfolio under 1) total ESG flows and 2) labelled ESG mutual fund flows. The second set of columns uses quarterly GMB (green-minus-brown) factor returns PST (2022). In specification (3) I instrument for the ESG flow by its lag F_{t-1}^{ESG} as in PST (2022). The first stage t-Statistic is 7.6. I only report IV results for ESG mutual fund flows, as the relevance condition does not hold for total ESG flows. For all specifications except for (3), I use rolling 4-quarter average flows. T-statistics are reported in parentheses. Significance at the 90, 95 and 99% confidence levels is indicated by *,**,*** respectively.

	ESG Return τ_t		GMB Factor Returns (PST, 2022)		
	(1)	(2)	(1)	(2)	(3)
const	-0.00	-0.00	0.02	0.02	0.02
	(-1.00)	(-0.21)	(2.17)	(2.66)	(2.15)
Total ESG Flow	2.47***		1.46		
	(3.45)		(1.46)		
ESG Mutual Fund Flow		47.65***		6.01	
		(3.40)		(0.08)	
IV (lagged Flow)					15.46
					(0.24)
R^2	0.29	0.29	0.01	0.00	-

A Stock-Specific ESG Flows and Returns

While the flows into individual stocks within the ESG portfolio are not directly observable, they can be approximated by aggregating the flow-driven trades of all investors. Total flow-driven trades in stock n are given by

$$\Delta d_{t+1,n} = \sum_{i=1}^{I} Q_{t,n}^{i} f_{t+1}^{i} \tag{2.15}$$

where I includes mutual funds and other 13F institutions (banks, pension funds, insurance companies etc.). f_{t+1}^i is the flow into fund i expressed in percent (i.e. dollar flow relative to lagged assets under management). $Q_{t,n}^i$ the fund's lagged ownership in stock n also expressed in percent because shares outstanding are normalized to 1. Precise data on flow-driven trades are only available for mutual funds. Because SEC 13F forms are filed at the management company level, flows towards ESG funds *within* a manager are difficult to capture.[33] Appendix B provides a detailed description on how to resolve this issue and approximate flow-driven trades for *all* 13F institutions. Note that because $Q_{t,n}^i$ are ownership shares, $\Delta d_{t+1,n}$ can be interpreted as demand shocks in percent relative to shares outstanding. Thus a flow-

[33] For example, a $1 billion exogenous flow from the Vanguard S&P500 ETF to the Vanguard FTSE Social Index Fund would only show up as a demand shock for greener stocks in Vanguard's aggregate share holdings while leaving its total assets under management unchanged.

driven demand shock of $\Delta d_{t+1,n} = 0.01$ implies a 1% increase in the demand for the stock. If flows affect the cross-section of ESG returns, then $\Delta d_{t,n}$ should be significantly related to the cross-section of ESG returns. Let $\Delta p_{t+1,n}$ denote the quarterly return on stock n. I compute abnormal returns $\alpha_{t+1,n}$ by cross-sectionally orthogonalizing returns with respect to market beta, log market equity, log market-to-book ratio, profitability and investment.[34] I then compute the cumulative flow-driven demand Δd_n and abnormal returns for every stock Δp_n by summing over the sample period from 2016 to 2021.

B Price Pressure in the Cross-Section of ESG stocks

The structural model predicts that flow-driven price pressure is given by multiplying the vector of flow-driven demand shocks Δd_t by the multiplier matrix $\mathscr{M}_t$. Thus, the model implies a 1:1 mapping between $\mathscr{M}_t \Delta d_t$ and the cross-section of realized ESG returns. Panel (a) of Figure 2.4 plots the abnormal returns on all green stocks along with the flow-driven price pressure. I fit a linear regression through the scatter points, which shows that the cross-section of green returns is significantly related to flow-driven price pressure (t-Statistic > 7). Furthermore, both value- and equal-weighted regression lines are close to the diagonal. This is strong evidence in favor of the overall magnitude of the demand elasticities obtained from holdings data.[35] On average, the multiplier matrix $\mathscr{M}_t$ correctly maps the cross-section of demand shocks into the realized return space. Panel (b) of Figure 2.4 plots the price pressure ($\mathscr{M}\Delta d$) for all green stocks against their abnormal return α_n. Once again, the cross-section of abnormal ESG returns is significantly related to price pressure with a t-statistic of 5.25. We furthermore cannot reject the null hypothesis that the slopes of the fitted lines are different from the diagonal.

C Testing different Elasticity Estimates

The flow-driven demand shocks allow for a more granular test of the price impact implied by demand-based asset pricing models. Note, that equation (2.7) is a first-order approximation to the demand shock in a large class of models, including the Demand System Approach to Asset Pricing by KY (2019).[36] We can test different elasticity estimates by comparing the cross-section of flow-driven ESG returns to the price pressure obtained from different multipliers $\mathscr{M}_t$. Panel (a) of Table 2.6 compares the regression slope and R^2 across different multipliers. The first column regresses the cross-section of abnormal ESG returns α_n onto the raw flow-driven demand shocks Δd_n, which implies a diagonal multiplier matrix equal to the identity matrix. The slope coefficient is 1.13 with a t-statistic of 6.84. This implies that a 1% demand shock for a green stock increases its price by 1.13%. The average multiplier across all green

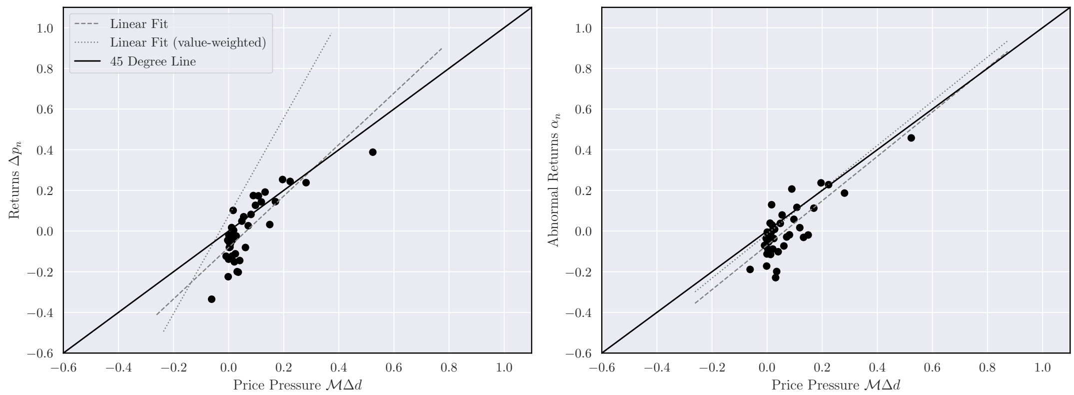
[34] Formally, I run the quarterly cross-sectional regressions of log returns $r_{t+1,n}$ onto lagged characteristics $X^k_{t,n}$ and extract the residual: $r_{t+1,n} = \beta^0_t + \sum_{k=1}^K \beta^k_t X^k_{t,n} + \Delta p_{t+1,n}$.

[35] If true multipliers were significantly higher than the model-implied estimates (i.e. investors respond more strongly to price changes than implied by the model) then the slope of the regression line would be much steeper.

[36] See van der Beck (2022) for details.

Figure 2.4: **ESG Demand in the Cross-Section of ESG Returns**
The figure reports binned scatter plots of the cross-section of ESG returns against flow-driven price pressure $\mathcal{M}_t \Delta d_{t,n}$. For each stock, I compute the cumulative price pressure and returns from 2016 to 2022. Panel (a) plots price pressure against raw cumulative returns Δp_n. Panel (b) plots price pressure against abnormal returns α_n obtained from cross-sectional regressions of returns onto known predictors.

(a) Raw Returns | (b) Abnormal Returns



stocks obtained from the structural model is 1.11, which is strikingly close to the reduced-form estimate. Raw flow-driven demand furthermore explains around 4% of the cross-section of ESG returns. The second column regresses the cross-section of abnormal returns onto demand shocks scaled by the multiplier matrix as in Figure 2.4. The explained variation of the cross-section of ESG returns rises to 5%. This does not necessarily imply that the additional information contained in the stock-specific and cross-elasticities is small. It rather confirms that the cross-section of individual stock returns is driven by unobservable latent demand shocks unrelated to flow-driven demand (see KY (2019)). The third column uses the elasticity matrix obtained using the methodology in KY (2019). Recall that they identify elasticities using portfolio holdings in levels, whereas this paper identifies elasticities from quarterly trades (i.e. changes in portfolios). The regression slope drops to 0.49 (t-Statistic of 7.22) which implies that the price pressure estimated from holdings in levels is slightly too large. This may be owed to the endogeneity problem of identifying elasticities from holdings as opposed to trades.[37]

Lastly, I provide a simple test of whether the stock-specific multipliers, i.e. the diagonal elements in $\mathcal{M}$ contain additional information about price pressure beyond traditional characteristics. To this end, let $\frac{\alpha_n}{\Delta d_n}$ denote a primitive measure of stock-specific price pressure. It is the abnormal return on stock n from 2016 to 2022 divided by the cumulative flow-driven demand Δd_n. I regress $\frac{\alpha_n}{\Delta d_n}$ in the cross-section onto the stock-specific multipliers, $\mathcal{M}_n$, controlling for log market equity and market beta. Panel (b) of Table 2.6 reports the estimated coefficients. The price impact of flow-driven demand in cross-section of green stocks is significantly larger for smaller stocks. More importantly, it is significantly positively related to the stock-specific price impact implied by the structural model (with a t-Statistic of 3.6).

[37] See van der Beck (2022) for details.

Table 2.6: **The Cross-Section of ESG Returns and different Elasticity Measures**
Table (a) reports the slope coefficient β_1 and R^2 of regressions of the following form

$$\alpha_n = \beta_0 + \beta_1 \text{Pressure}_n + \epsilon_n$$

where Pressure_n is 1) raw demand Δd_n, 2) demand multiplied by the multiplier matrix estimated in this paper, and 3) demand multiplied by the multiplier matrix estimated in KY (2019). Panel (b) plots the coefficient estimates of a panel regression of price impact $\frac{\alpha_n}{\Delta d_n}$ onto the diagonal elements of the elasticity matrix, controlling for log market equity and beta. T-statistics are reported in parentheses.

(a) Comparing Multiplier Estimates

	Raw Demand Δd_n	Demand × Multiplier $\mathscr{M}\Delta d_n$	Demand × KY-Multiplier $\mathscr{M}_{KY}\Delta d_n$
$\Delta\alpha_n$	1.13	1.09	0.44
	(6.84)	(7.28)	(5.90)
R^2	0.04	0.05	0.03

(b) Price Pressure Cross-Section

	Price Impact $\alpha_n/\Delta d_n$
const	1.809
	(1.418)
diag $\mathscr{M}$	2.813***
	(3.598)
Log ME	-0.211***
	(-4.488)
Beta	0.154
	(1.313)

2.6 Applications and Robustness Tests

A ESG Index Inclusion

The structural approach presented in this paper allows to circumvent the issue that ESG flows and returns are jointly endogenous. It is nevertheless reassuring if the structural estimates are at least to some extent backed by simple reduced-form evidence, such as demand shocks from ESG index inclusions. A well-known ESG index is the FTSE USA 4 Good Index (henceforth 4G Index). Berk and van Binsbergen (2022) use a stock's membership in the 4G Index as a proxy for aggregate ESG demand and find that there are no price effects associated with inclusion in the index. They conclude that impact investing does not affect firms' cost of capital. However, it is unclear how much money is actually flowing into the stocks added to the index. In other words, are the assets indexed to the 4G Index large enough to generate meaningful demand shocks based on its reconstitution?

To further investigate this, I construct mutual fund demand $\Delta q_{t,n}^{MF}$ as the change in ownership by mutual funds for every stock n and quarter t. Table 2.7 reports regressions in the style of Berk and van Binsbergen (2022). $\Delta I_{t,n}^{4G}$ is a variable equal to 1 in the quarter of inclusion in the 4G Index, -1 in the quarter of exclusion, and 0 otherwise. I first regress index flow onto $\Delta I_{t,n}^{4G}$ including the controls used in Berk and van Binsbergen (2022).[38] Addition to 4G Index is

[38]Because additions and deletions are encoded as 1 and -1 respectively, I refer to all index reconstitutions as additions.

associated with a significant increase in total mutual fund ownership of 1.3 percent. In other words, when a stock is included in the 4G Index, mutual funds contemporaneously purchase 1.3 percent of the stock's shares outstanding (on average). This suggests that the 4G Index is sufficiently widely followed such that reconstitutions cause meaningful shocks to index investor demand. The third column of the table replicates the specification in Berk and van Binsbergen (2022) at a quarterly frequency by regressing quarterly stock returns onto $\Delta I^{4G}_{t,n}$. As in their study, the coefficient is insignificant and very small, suggesting that the ESG flows do not generate meaningful price pressure. However, for only 56% of all index reconstitutions, mutual fund flow $\Delta q^{MF}_{t,n}$ has the same sign as $\Delta I^{4G}_{t,n}$.[39] In order to identify relevant (i.e. widely followed) reconstitution events, I use ownership changes of index trackers. I interact the 4G index reconstitutions with a dummy variable, $\mathbb{1}_{\text{Demand}}$, equal to 1 if the demand by index trackers has the same sign as the reconstitution.[40] The coefficient on the interaction term is large and statistically significant. This suggests that ESG index inclusion has a strong effect on prices, as long as mutual funds actually purchase the stock when it is included. In other words, conditional on inclusion in the 4G index, the demand by index trackers has a large impact on the prices of green firms. Furthermore, the implied price impact is in line with the multiplier obtained from the structural model. In the quarter of inclusion in the 4G index, the stocks followed by index trackers receive a $3.2 - 0.7 = 3.13\%$ demand shock by mutual funds and experience $11.5 - 6.2 = 5.3\%$ higher returns, which implies an ESG demand multiplier of $\frac{5.3}{3.2} = 1.69$. The dummy variable $\mathbb{1}_{\text{Demand}}$ measures the demand by *index-tracking* funds and should therefore not contain not contemporaneous return-chasing behaviour. Nevertheless, endogeneity concerns remain because index-trackers often focus on the largest or most liquid stocks within the index, which may precisely be the ones that had high returns. I therefore construct an alternative dummy variable that is agnostic to the sign of the demand by index trackers. To this end, I define index turnover as the total trading volume by index-tracking mutual funds. I then interact the 4G index reconstitutions with a dummy variable, $\mathbb{1}_{\text{Turnover}}$, equal to 1 if the turnover by index trackers is over two standard deviations away from its stock-specific mean.[41] Column (3) of both panels report the results. High index-turnover stocks have a $7.4 + 0.5 = 7.9\%$ higher mutual fund demand and experience $6.8 + 0.1 = 6.9\%$ higher returns during the inclusion quarter in the 4G Index. This implies a multiplier of $\frac{6.9}{7.9} = 0.87$.

[39] From 2012 to 2021 and using quarterly data, I obtain 342 reconstitution events of the 4G index, conditional on the stock already being in the FTSE USA Index. For 192/342=56% out of these events, the aggregate ownership change of mutual funds has the same sign as the reconstitution.

[40] In particular, $\mathbb{1}_{\text{Demand}}$ equal to 1 if the sign of index fund flow during the reconstitution quarter is the same as the sign of the reconstitution $\Delta I^{4G}_{t,n}$. Index fund flow is defined as the change in ownership by index trackers. To identify index funds, I use the label 'Pure Index Fund' provided by the CRSP mutual fund database, which are mutual funds with an index-fund flag equal to 'D'.

[41] More precisely, index turnover is given by $\text{IT}_{t,n} = \frac{\sum_{i \in I^x} |\Delta Q^i_{t,n}|}{\sum_{i \in I^x} Q^i_{t-1,n}}$ where $I^x \subset I$ is the subset of index-tracking mutual funds. The unconditional time-series mean and standard deviation are μ^{IT}_n and σ^{IT}_n. The index turnover indicator, $\mathbb{1}_{\text{Turnover}}$, is equal to 1 if $\text{IT}_{t,n} - \mu^{\text{IT}}_n > 2\sigma^{\text{IT}}_n$.

Table 2.7: **How much money is following the FTSE 4 Good Index?**
The table reports different regressions in the style of Berk and van Binsbergen (2022). $\Delta I^{4G}_{t,n}$ is equal to 1 in the quarter of inclusion in the FTSE 4 Good Index, -1 in the quarter of exclusion, and 0 otherwise. $I^{4G}_{t,n}$ is a dummy equal to 1 in all quarters after inclusion. $\Delta I_{t,n}$ and $I_{t,n}$ are defined equivalently, but for the FTSE USA index. The stocks in the FTSE 4 Good Index are a strict subset of the stocks in the FTSE USA index. $\mathbb{1}_{\text{Demand}}$ is a dummy equal to 1 if aggregate purchases by index-tracking mutual funds have the same sign as the $\Delta I^{4G}_{t,n}$. $\mathbb{1}_{\text{Turnover}}$ is a dummy equal to 1, if the turnover by index-tracking mutual funds is more than 2 standard deviation away from its stock-specific mean. The first two columns use quarterly mutual fund flow as a dependent variable. Columns three and four use quarterly stock returns. T-statistics are in parentheses. Standard errors are double clustered at the stock and year-month level. Significance at the 90, 95 and 99% confidence levels is indicated by *,**,*** respectively.

	Mutual Fund Flow $\Delta q^{MF}_{t,n}$			Quarterly Returns $\Delta p_{t,n}$		
	(1)	(2)	(3)	(1)	(2)	(3)
const	0.002	0.002	0.002**	0.047**	0.047**	0.047**
	(1.64)	(1.64)	(1.64)	(2.43)	(2.43)	(2.43)
$I_{n,t}$	-0.001	-0.001	-0.001	-0	-0.001	-0
	(-0.93)	(-1.01)	(-0.81)	(-0.06)	(-0.1)	(-0.04)
$I^{4G}_{n,t}$	-0.001	-0.001	-0.001	-0.005	-0.005	-0.005
	(-1.2)	(-1.18)	(-1.41)	(-0.8)	(-0.78)	(-0.82)
$\Delta I_{n,t}$	-0.001	-0.001	-0.002	-0.001	-0.001	-0.001
	(-0.51)	(-0.54)	(-0.82)	(-0.03)	(-0.04)	(-0.06)
$\Delta I^{4G}_{n,t}$	0.013***	-0.007*	0.005***	0.009	-0.062*	0.001
	(4.5)	(-1.91)	(2.78)	(0.29)	(-1.75)	(0.05)
$\Delta I^{4G}_{n,t} \times \mathbb{1}_{\text{Demand}}$		0.032***			0.115**	
		(4.52)			(2.21)	
$\Delta I^{4G}_{n,t} \times \mathbb{1}_{\text{Turnover}}$			0.074***			0.068**
			(10.06)			(2.38)
R^2	0.00	0.00	0.00	0.00	0.00	0.00
Observations	125263	125263	125263	125263	125263	125263

B Mandate-Driven Portfolio Reconstitutions

Mutual fund purchases based on additions and deletions from the 4G Index represent a small set of potentially exogenous ESG demand shocks. In this section, I generalize the idea of ESG index inclusion to construct a larger set of exogenous ESG demand shocks. In order to disentangle non-fundamental from fundamental ESG demand, it will be useful to define two kinds of demand shocks: Intensive and extensive. Intensive demand shocks are changes in the shares held by an investor that do not originate from or result in zero holdings. Extensive demand shocks are portfolio additions and deletions, i.e. changes in shares held originating from or resulting in zero holdings. A key difference between the two is that extensive demand shocks likely contain an exogenous (non-fundamental) component that is related to the investment mandate of the fund. For example, an ESG investor may include a stock in her portfolio once the company's Co2 Emissions fall below the industry median. Similarly, a value fund includes a stock if it falls in the top quintile of book to market ratios. While not all of the extensive demand shock is non-fundamental (Co2 emissions dropped because of a change in production which affects profits) at least part of it is driven by the fund's exogenous portfolio constraint: "Buy the bottom 50% of Co2 emitters". Let $\Delta Q^{\perp}_{t,n}$ denote the total amount of shares purchased due to specific ESG mandates and portfolio constraints. Because $\Delta Q^{\perp}_{t,n}$ is orthogonal to fundamental news, a significant relationship with contemporaneous returns $r_{t,n}$ would confirm that non-fundamental ESG demand affects prices. Appendix Section B shows how to construct $\Delta Q^{\perp}_{t,n}$ from mutual funds' extensive and intensive trades.

In order to test whether non-fundamental ESG demand impacts prices, I estimate panel regressions of quarterly returns onto $\Delta Q^{\perp}_{t,n}$ controlling for known return predictors such as market beta, size, value, profitability, and investment. Table 2.8 reports the estimated coefficients on $\Delta Q^{\perp}_{t,n}$ for different specifications. The coefficient on $\Delta Q^{\perp}_{t,n}$ is highly statistically significant with a t-Statistic of 11.02. Note, that $\Delta Q^{\perp}_{t,n}$ only captures the exogenous demand shocks of labelled ESG mutual funds, which represent a subset of all ESG investors. If we scale $\Delta Q^{\perp}_{t,n}$ by the inverse market share of labelled ESG funds relative to total ESG assets, we can identify the structural parameter linking exogenous ESG demand and prices.[42] The coefficient on scaled $\Delta Q^{\perp}_{t,n}$ is 0.95, which implies that when ESG funds purchase 1% of a company's shares outstanding, the price increases by roughly 0.95%. The implied multiplier is 0.95 which is close to the estimate from the structural model of 1.11. I also sort $\Delta Q^{\perp}_{t,n}$ into quartiles by absolute values and assign dummy variables equal to 1 if $\Delta Q^{\perp}_{t,n}$ in the respective quartile is positive and -1 if $\Delta Q^{\perp}_{t,n}$ in the respective quartile is negative. The results show that stocks with higher mandate-driven ESG demand experience stronger price pressure. The coefficient on $\Delta Q^{\perp}_{t,n}$ is significant across all specifications. Thus ESG investors' trades that are driven by portfolio constraints and investment mandates have a significant impact on prices. Furthermore, the magnitudes are consistent with the elasticities estimated from holdings data.

C Impact-Investing at the Fund Level

The interaction between the multiplier matrix $\mathcal{M}_t$ and fund-specific deviations from the market portfolio allows for assessing the efficacy of impact-investing at the fund level. A fund's impact is driven by its deviation from the market portfolio and by the extent to which the

[42] Formally, total $\Delta Q^{\perp}_{t,n} = S^{ESG}_t \Delta Q^{\perp,\text{Total}}_{t,n}$ where $Q^{\perp,\text{Total}}_{t,n}$ is total mandate-driven ESG demand and S^{ESG}_t is the market share of labelled ESG mutual funds. Thus $Q^{\perp,\text{Total}}_{t,n} = \frac{1}{S^{ESG}_t} \Delta Q^{\perp}_{t,n}$.

Table 2.8: **Price Impact of Non-Fundamental ESG Demand**
The table reports the results of panel regressions of quarterly returns onto non-fundamental ESG demand $\Delta Q^{\perp}_{t,n}$. Specification (1) uses the raw $\Delta Q^{\perp}_{t,n}$, which are mandate-driven portfolio additions by labelled ESG mutual funds. Specification (2) scales $\Delta Q^{\perp}_{t,n}$ by the inverse market share of labelled ESG funds relative to total ESG assets. Specification (3) includes changes in fundamentals as additional controls. Specification (4) splits $\Delta Q^{\perp}_{t,n}$ into quartile dummies. Robust t-Statistics are reported in parentheses. Significance at the 90, 95 and 99% confidence levels is indicated by *,**,*** respectively.

	Quarterly Returns			
	(1)	(2)	(3)	(4)
$\Delta Q^{\perp}$	58.95***			
	(11.02)			
$\Delta Q^{\perp}$ Scaled		0.95***	0.88***	
		(10.44)	(9.88)	
$\mathbb{1}(\Delta Q^{\perp}$ Quartile 1)				-0.008***
				(-3.240)
$\mathbb{1}(\Delta Q^{\perp}$ Quartile 2)				0.001
				(0.250)
$\mathbb{1}(\Delta Q^{\perp}$ Quartile 3)				0.001
				(0.330)
$\mathbb{1}(\Delta Q^{\perp}$ Quartile 4)				0.026***
				(13.150)
Fundamental Controls	Yes	Yes	Yes	Yes
Changes in Fundamentals	No	No	Yes	No
Time FE	Yes	Yes	Yes	Yes

deviations are concentrated towards inelastic stocks.[43] A fund's ability to affect green firms' cost of capital is strongly limited, if it overweights stocks that are held by elastic investors and by investors who respond by substituting towards brown stocks. Green stocks that are associated with a high multiplier are best suited for impact investing as flows induce a large realized return and hence a lower cost of capital. Also, note that $\mathcal{M}_t$ is an $N \times N$ matrix that accounts for flow-driven spillover effects to all stocks. If the market accommodates green demand primarily by substituting towards other green stocks, then $\mathcal{M}_{t,gg}$ is high, causing an amplified relative price impact. Table 2.9 reports the impact of a \$1 flow from the market portfolio towards the largest ESG mutual funds averaged over the past 5 years. There is great heterogeneity in the funds' impact on green stocks. A \$1 flow to the Calvert Social Investment Fund boosts the aggregate value of green stocks by \$0.82. In contrast, the same flow towards the Vanguard FTSE Social Index Fund raises the value of green stocks by only \$0.39. Furthermore, many sustainable funds unintentionally boost the value of fossil fuel companies. A \$1 flow

[43] Formally, fund i's impact is given by $\mathcal{M}_t w^i_{t,n} - w^{MF}_{t,n}$. These are the cross-sectional price changes due to a 1 dollar from from the aggregate mutual fund portfolio towards i.

Table 2.9: **Flow Impact at the Fund Level**
The table reports the impact of a $1 flow towards some of the largest ESG mutual funds in the US. I compute the impact at every quarter and then average across quarters from 2016 to 2021. I report the impact on green stocks for which $\tau_{t,n} > 0$, as well as fossil fuel and sin stocks. The second column reports the funds' active deviation from the S&P 500 computed as $\frac{1}{2}\sum_n |w^i_{t,n} - w^{SP}_{t,n}|$.

		Impact of 1$ Flow onto...		
	Deviation from S&P500	Green Stocks	Fossil Fuel Stocks	Sin Stocks
TIAA-CREF Funds: Social Choice Equity Fund	0.546	0.272	0.013	-0.008
Calvert Social Investment Fund	0.819	0.613	-0.006	-0.010
Putnam New Opportunities Fund	0.762	0.173	-0.052	-0.027
Vanguard FTSE Social Index Fund	0.391	0.098	-0.047	-0.030
Calvert Social Index Fund	0.339	0.097	-0.041	-0.021
Virtus Small-Cap Sustainable Growth Fund	0.971	-0.105	0.052	0.016
iShares FTSE KLD 400 Social Index Fund	0.594	0.313	-0.000	-0.015
Brown Advisory Winslow Sustainability Fund	0.833	0.432	-0.006	-0.009
iShares MSCI USA ESG ETF	0.515	-0.043	0.030	0.007

towards the iShares MSCI USA ESG ETF increases the aggregate value of fossil fuel companies by 0.03$ and decreases the value of green stocks by 0.04$. On the one hand, this heterogeneity is owed to the fact that there is no objective measure of a fund's *true* sustainability. Asset managers use different sustainability metrics, which often diverge substantially (see Berg et al. (2019) and Berg et al. (2021)). On the other hand, funds differ strongly in their deviation from the market portfolio. Some funds, such as the Vanguard FTSE Social Index Fund or the iShares MSCI USA ESG ETF, deviate very little from S&P500 index weights and hence primarily serve as a way for investors to feel good about themselves without having a *true* impact. Surprisingly, flows towards many sustainable funds raise the aggregate valuation of fossil fuel companies. Even though an ESG fund may underweight an industry as a whole, by tilting towards more inelastic stocks it can positively affect the aggregate valuation of that industry. Similarly, if the fund tilts towards stocks that have high cross-elasticities with underweighted stocks it unintentionally boosts the valuation of the wrong companies. Overall, Table 2.9 emphasizes that while sustainable flows do impact firms' realized returns and cost of capital, the choice of the appropriate fund is crucial to affect change in the preferred direction.

2.7 Conclusion

This paper investigates the extent to which the realized returns from ESG investing are owed to price-pressure arising from flows towards sustainable funds. Flow-driven price pressure is the product of sustainable funds' deviation from the market portfolio and the market's elasticity of substitution between stocks. I find that every dollar flowing from the market portfolio towards the ESG portfolio increases the aggregate value of green firms by \$0.4. Further, ESG funds would have likely underperformed the market in the absence of flow-driven price pressure on green stocks. Thus, one should be careful when using the *realized* outperformance of sustainable investments in recent years to judge their *expected* outperformance going forward. While the low aggregate elasticity of substitution is worrying for the overall stability and efficiency of equity markets, it supports the effectiveness of impact investing. Flows towards green funds that invest in cross-sectionally inelastic stocks substantially reduce the cost of capital of the firms in the funds' portfolios. As the framework allows quantifying the effect of flows on green firms' cost of capital, it enables differentiating sustainable funds by an objective *real-impact* criterion.

The large impact of flows on realized ESG returns has important consequences for *expected* ESG returns going forward. Assessing the extent to which *expected* returns are affected by demand pressures is non-trivial as it depends on the *expected* flows into ESG funds. If ESG funds continue to receive inflows then the prices of green firms will further increase causing positive realized returns in the future. The reduction in short-term expected returns due to flow-induced price pressure is therefore small. If, however, ESG inflows unexpectedly revert, the realized future return may be strongly negative. The question, whether ESG funds will receive outflows in the future ultimately depends on whether ESG flows are performance- or taste-based. It is likely that at least some flows to ESG funds are driven by past performance rather than *true* shifts in green preferences. Even if none of the flows to ESG funds are performance-driven, green preferences fluctuate over time and may well decline during bad economic times. Importantly, *expected* ESG returns going forward also depend on the transitory versus permanent nature of past demand shocks. van der Beck (2022) shows, that investors become more elastic in the long run. The impact of demand shocks on equilibrium prices therefore partly reverts over time. In other words, over a longer horizon investors substitute away from overpriced green stocks. As ESG funds move closer to the market portfolio, this effect may outweigh the price impact of continued ESG flows. Investigating the implications of demand shocks, elasticities, and arbitrage in a dynamic context is an important avenue for further research.

Lastly, the purpose of impact investing goes beyond (temporarily) boosting the stock prices of sustainable companies. Do sustainable firms capitalize on the rise of ESG investing by issuing new shares at elevated prices and undertaking green projects? Investigating the *real* effects of flow-driven price pressure by providing an explicit link between demand-based asset pricing and corporate finance opens up an exciting research agenda.

2.8 Proofs and Supplementary Material

A The ESG Portfolio

Robustness of the ESG portfolio

The ESG portfolio is constructed using ESG mutual funds' portfolio holdings. To this end, I identify a large set of ESG mutual funds via their fund name as reported by CRSP. A mutual fund is an ESG fund if its name contains at least one (or any abbreviation) of a list of sustainability keywords :*Environment, social, governance , green, sustainable, responsible, SRI, ESG, climate, clean, carbon, impact, fair, gender, solar, earth, renewable, screen, ethical, conscious, CSR, thematic*. The total list of keywords is much larger. For brevity, this list excludes all keywords that are not actually used in funds' names. Figure 2.10 plots the 30 largest ESG funds and their assets under management as of December 2021.

Table 2.10: **Largest 30 ESG Funds**
The table the largest 30 ESG funds and their assets under management identified by the list of sustainability keywords. Assets under management are reported in billion USD.

Fund Name	Assets	Fund Name	Assets
iShares ESG Aware MSCI USA ETF	25.70	iShares MSCI KLD 400 Social ETF	4.20
Vanguard FTSE Social Index Fund	16.79	Xtrackers MSCI USA ESG Leaders Equity ETF	4.14
TIAA-CREF Social Choice Equity Fund	7.75	Sustainable Equity Fund	3.86
iShares ESG Aware MSCI EAFE ETF	7.62	International Sustainability Core 1 Portfolio	3.47
Brown Advisory Sustainable Growth Fund	7.38	CCM Community Impact Bond Fund	3.40
Core Impact Bond Fund	7.27	Vanguard ESG International Stock ETF	3.17
Putnam Sustainable Leaders Fund	6.82	Calvert Small Cap Fund	3.06
Calvert Impact Fund	6.75	Invesco Floating Rate ESG Fund	2.86
Vanguard ESG US Stock ETF	6.50	FT Clean Edge Green Energy Index Fund	2.82
iShares ESG Aware MSCI EM ETF	6.22	Pax Global Environmental Markets Fund	2.74
US Sustainability Core 1 Portfolio	5.86	Invesco Solar ETF	2.73
iShares Global Clean Energy ETF	5.61	AB Sustainable Global Thematic Fund	2.65
Calvert US Large-Cap Core Resp. Index Fund	5.26	Pax Sustainable Allocation Fund	2.62
iShares MSCI USA ESG Select ETF	4.82	Calvert Bond Fund	2.49
iShares ESG MSCI USA Leaders ETF	4.31	PIMCO Total Return ESG Fund	2.46

Note, that the ESG portfolio w_t^{ESG} is scale-invariant and does not depend on the number of identified ESG funds. Its representativeness therefore only depends on whether the subset of ESG funds identified via the list of keywords is represenative of the total ESG fund population. In other words, how stable is w_t^{ESG} for different samples of ESG funds. At every quarter, I sort the sample of ESG funds by their assets under management and split the sample in two groups based on whether a fund has an odd or even rank. I then aggregate the holdings for the two groups and compute two representative ESG portfolios $w_t^{ESG,even}$ and $w_t^{ESG,odd}$. The two portfolios are therefore computed using two different (non-overlapping) subsets of funds. I also define two measures of greenness τ_t^{even} and τ_t^{odd} as the deviation of $w_t^{ESG,even}$ and $w_t^{ESG,odd}$ from the aggregate mutual fund portfolio w_t^{MF}.[44]. Figure 2.5 plots the quarterly cross-sectional correlation of the two ESG portfolios and the two taste measures. I plot both raw (i.e. equal-weighted) correlations, and market cap-weighted correlations.

Figure 2.5: **Representativeness of the ESG Portfolio**
Panel (a) plots quarterly cross-sectional correlations between $w_t^{ESG,even}$ and $w_t^{ESG,odd}$. Panel (b) plots the quarterly cross-sectional correlations between τ_t^{even} and τ_t^{odd}, which are deviations of the ESG portfolios from the aggregate mutual fund portfolio w_t^{ESG}. I compute both equal-weighted and market cap-weighed correlations and plot 3-month rolling averages of the cross-sectional correlation coefficients.

(a) Correlation $w_t^{ESG,even}$ and $w_t^{ESG,odd}$ (b) Correlation τ_t^{even} and τ_t^{odd}

The two ESG portfolios are highly correlated with correlations above 90% for the later part of the sample. This correlation is not just by driven the common tilt towards the aggregate mutual fund portfolio. The ESG portfolio's deviations from the aggregate mutual fund portfolio, τ_t^{even} and τ_t^{odd}, are also highly correlated with an average correlation above 50%. Market-cap weighted correlations are slightly higher implying that there is stronger agreement among ESG funds for larger stocks.

Investor Preference for ESG Labels

In light of the large flows to sustainable funds in recent years, a natural question that arises is whether including an ESG keyword in the fund title leads to increased inflows. In other words, can fund managers effectively *buy* additional flows by simply changing their fund's name?

Let $\mathbb{1}_{ESG,t}^{i}$ denote a dummy variable equal to 1 if fund i has an ESG keyword in its name at

[44]Formally $\tau_t^{even} = w_t^{ESG,even} - w_t^{MF}$ and $\tau_t^{odd} = w_t^{ESG,odd} - w_t^{MF}$

date t. As a first preliminary test, I regress the panel of quarterly aggregated flows onto $\mathbb{1}^i_{ESG,t}$ controlling for lagged flows, fund size, fund performance, portfolio tilts and factor exposures. Panel (a) of Table 2.11 reports the estimated coefficient on the ESG dummy across different specifications.

Table 2.11: **ESG Labels and Flows**

The table reports the results to panel regressions of quarterly flows onto ESG indicators from 2010 to 2020. Panel (a) reports the coefficient on the ESG dummy equal to 1 if fund i has an ESG keyword in its name as of time t. The first column reports the specification without any controls except quarter fixed effects. The second column includes fund-level controls given by log assets under management, annual return, Sharpe ratio, Fama and French 3-Factor alpha, and flows lagged up to 9 quarters. The third column includes portfolio-level controls given by exposure to momentum, value and size factors as well as characteristic scores for momentum, value, size and greenness. Panel (b) reports the coefficients to the three dummy variables indicating whether a fund is an ESG fund at some point in the sample ($\mathbb{1}^i_{ESG}$), whether it changed its name to an ESG title at some point in the sample, and whether a previously non-ESG fund added an ESG keyword to its title ($\delta^i_{ESG,t}$). Standard errors across specifications are clustered at the fund level. Significance at the 99%, 95% and 90% level is indicated with ***,**,* respectively.

(a) ESG Fund Indicator $\mathbb{1}^i_{ESG,t}$

	Flows f^i_{t+1}		
	(1)	(2)	(3)
$\mathbb{1}^i_{ESG,t}$	0.041***	0.023***	0.020***
	(0.005)	(0.003)	(0.003)
Time FE	Yes	Yes	Yes
Fund Controls	No	Yes	Yes
Ptfl. Controls	No	No	Yes

(b) Name Change $\delta^i_{ESG,t}$

	Flows f^i_{t+1}		
	(1)	(2)	(3)
$\mathbb{1}^i_{ESG}$	0.047***	0.025***	0.022***
	(0.006)	(0.004)	(0.004)
$\mathbb{1}^i_{\text{treat}}$	-0.074***	-0.038***	-0.032***
	(0.008)	(0.004)	(0.005)
$\delta^i_{ESG,t}$	0.021**	0.021***	0.018**
	(0.009)	(0.006)	(0.008)
Time FE	Yes	Yes	Yes
Fund Controls	No	Yes	Yes
Ptfl. Controls	No	No	Yes

The estimates reveal that having an ESG keyword in the title leads significantly larger quarterly flows of 2%. Given that average quarterly flows are of the same magnitude, the flow gains from being regarded as an ESG fund are extremely large. The flow gain remains large and statistically significant at any reasonable confidence levels despite controlling for various fund-level characteristics including the lagged fund return, Sharpe ratio, Fama and French 3-Factor alpha, fund size (log assets under managament) and lagged flows up to 8 quarters. I also control for portfolio exposures to momentum, value and size obtained from regressing monthly fund returns onto factor returns. Lastly, I control for fund-level characteristic scores as in Lettau et al. (2018), which are portfolio-weighted averages of the stock characteristics.[45]

Nevertheless, it is possible (although unlikely) that ESG funds differ from other funds along some other dimension not captured by directly observable fund characteristics or common

[45] For every fund, I compute scores for greenness, value, size and momentum.

risk exposures and portfolio tilts. To address remaining endogeneity concerns, I use funds' name changes as exogenous variation in ESG titles. As mentioned in the previous section, 88 out of the 551 identified ESG funds have changed their name at some point between 2010 and 2020 by including an ESG keyword in the title. The name changes can be used akin to a difference-in-difference estimator in order to control for unobservable flow heterogeneity at the fund level. More formally, note that one can decompose the ESG dummy $\mathbb{1}^i_{ESG,t}$ into three sub-variables: i) A dummy $\mathbb{1}^i_{ESG}$ equal to 1 if the fund had an ESG keyword in the title at some point between 2010 and 2020, ii) a treatment dummy $\mathbb{1}^i_{\text{treat}}$ equal to 1 if the fund switched to an ESG title at some point between 2010 and 2020, and iii) a time-varying indicator $\delta^i_{ESG,t}$ equal to 1 *after* a previously non-ESG fund added an ESG keyword to its title. Any flow heterogeneity from having an ESG keyword in the title that is driven by some unobservable fund-level fixed effect is captured by $\mathbb{1}^i_{ESG}$ and $\mathbb{1}^i_{\text{treat}}$, such that $\delta^i_{ESG,t}$ captures the pure effect of the name change. Panel (b) reports the estimates across different specifications. Despite having very few name changes in the sample, the coefficient on the name change indicator $\mathbb{1}^i_{ESG}$ is statistically significant and roughly equal to 2%. Thus controlling for various fund-level characteristics, changing the fund's name to include an ESG keyword boosts quarterly flows by 2%. Note, that the treatment dummy $\mathbb{1}^i_{\text{treat}}$ is significantly negative. Thus funds with strong outflows seem to have a greater incentive to include trending ESG keywords in their name, which significantly alleviates subsequent outflows. This is an interesting avenue for further research.

Perceived versus True Sustainability

Do sustainable mutual funds invest sustainably? As already suggested in table 2.1, the ESG portfolio w^{ESG}_t tilts over 50% of its assets away from the aggregate mutual fund portfolio w^{MF}_t. However, this does not imply that ESG funds (in aggregate) tilt towards *truly* sustainable stocks. The difficulty in answering the question about *true* sustainability lies in the lack of an objective definition. Particularly the social and governance component of ESG investing may strongly depend on personal preferences and ethical convictions. While the environmental component may be more easily objectifiable (e.g. via Co2 Emission data), it is still subject to large variations in preferences. For example, is the least polluting company among all fossil fuel companies a sustainable company? Analyzing, which companies are *truly* sustainable lies beyond the scope of this paper. I nevertheless assess whether the ESG portfolio's deviations from the market portfolio align with a set of sustainability characteristics. To this end, I estimate two regressions. The first is a panel OLS regression of $\tau_{t,n}$ onto the sustainability characteristics. The second is a probit regression of a greenness dummy $\mathbb{1}_{\tau>0}$ (which is equal to 1, if the stock is overweighted by the ESG portfolio) onto the same set of sustainability characteristics. As sustainability characteristics, I use Refinitiv ESG Scores, a Co2 emissions indicator, a sin stock dummy, a fossil fuel industry dummy and a Vanguard 4 Good dummy equal to 1, if the stock is in the Vanguard 4 good index. The Co2 emissions indicator is equal to 1 at time t, if the stock is in the highest decile of Co2 scope 1 emissions across all stocks in the sample. I furthermore control for log market equity, market beta and volatility in both specifications. Table 2.12 reports the results.

Table 2.12: **ESG tastes and sustainability characteristics**
The table reports the results of two regressions. The first is a panel regression including time fixed effects of $\tau_{t,n}$ onto different sustainability characteristics. The second is a probit regression of a greenness dummy $\mathbb{1}_{\tau>0}$ (which is equal to 1, if the stock is overweighted by the ESG portfolio) onto the same set of sustainability characteristics. The control variables in all specifications are log market equity, market beta and volatility. *,**,*** denote significance at the 90, 95 and 99% confidence level.

	Sustainability Characteristics						
	ESG Score	High Co2 Emissions	Sin Stock	Fossil Fuel	Vanguard 4 Good Index	Controls	R^2
Panel Regression $\tau_{t,n}$							
Coefficient	0.21***	-0.21***	-0.58***	-0.21***	0.51***	Yes	5.92%
t-stat	11.01	-16.28	-13.57	-11.80	39.05		
Probit Regression $\mathbb{1}_{\tau>0}$							
Coefficient	0.78***	-0.33***	-0.47***	-0.14***	0.46***	Yes	5.53%
t-stat	28.83	-23.20	-9.82	-5.39	41.88		

The coefficients on virtually all sustainability characteristics are highly significant with the right sign. The ESG portfolio tilts significantly towards stocks with high ESG scores as well as stocks that are in the Vanguard 4 Good Index. It significantly underweights sin stocks, stocks in the fossil fuel industry and high Co2 emitters. This is strong evidence, that ESG funds (on aggregate) do tilt towards what may be labelled as *objective* sustainability. Kim and Yoon (2022), Liang et al. (2021) and Gibson et al. (2022), on the other hand, show that investors who are part of the Principles for Responsible Investment initiative do not have better ESG scores. The opposing results underline the above-mentioned concerns that treating readily available scores by ESG ratings providers as *objective* or *true* sustainability is problematic. Recent evidence furthermore suggests, that ESG scores by ratings providers are inflated by greenwashing and empty sustainability claims (see Yang (2021) and Bams and van der Kroft (2022)).

B Measuring ESG Flows

Aggregate ESG Flows

Price pressure in aggregate ESG returns is driven by flows towards the ESG portfolio w_t^{ESG}. Total cumulative flows into labelled ESG mutual funds from Section 3.3 amount to roughly \$175 billion as of December 2021. However, the flows into labelled ESG do not include the (unobservable) ESG tilts of other mutual funds, large investment advisors, pension funds, banks, insurance companies, and other institutions. Unfortunately, precise data on flows are only available for mutual funds. 13F institutions report their holdings at the management company level. Thus flows towards ESG funds *within* a manager show up as active trades a^i_{t+1} instead of flow-driven trades $Q^i_t f^i_{t+1}$. To illustrate this point, consider the following simple example.

Example. Manager i manages two investment funds, an ordinary index fund and an ESG fund that overweights green stocks and underweights brown stocks. Between t and $t+1$ investors withdraw money from the index fund and invest it in the ESG fund provided by the same manager. Thus total flows f^i_{t+1} are 0, but the manager buys some green stocks ($\Delta Q_{t+1,\text{green}} > 0$) and sells some brown stocks ($\Delta Q_{t+1,\text{brown}} < 0$). In the aggregated 13F holdings, these trades only show up as active trades a^i_{t+1}, even though they are purely flow-driven.

In order to address this issue I propose decomposing 13F institutions' portfolios into different fund-level portfolios via a simple cross-sectional projection. For simplicity of notation, I am dropping the fund superscripts i. For every 13F-quarter pair, I am projecting the portfolio weights onto a set of $s = 1, \ldots, S$ managed portfolios (or individual funds)

$$\begin{aligned} \min_{\{\beta^s_t\}^S_{s=1}} \quad & \| w_{t,n} - \sum_{s=1}^{S} \beta^s_t w^s_{t,n} \|_2 \\ \text{s.t.} \quad & 0 \leq \beta^s_t \leq 1 \quad \forall s = 1, \ldots, S \end{aligned} \tag{2.16}$$

Thus β^s_t are the wealth-shares of individual funds belonging to institution i and $w^s_{t,n}$ their corresponding portfolios. As a set of managed portfolios $w^s_{t,n}$, I choose the equal-weighted portfolio $w^E_{t,n} = 1/N^i$, the market cap-weighted portfolio $w^{Mkt}_{t,n} = P_{t,n}/\sum_{n\in N^i} P_{t,n}$ and the ESG portfolio $w^{ESG}_{t,n}$. (2.16) is essentially equal to a constrained cross-sectional regression of portfolio weights $w_{t,n}$ onto a constant (the equal-weighted portfolio $w^E_{t,n}$) and characteristics (the other managed portfolios). The managed portfolios are constructed such that the weights sum to 1 across the institution's current holdings N^i. This implies rescaling the ESG portfolio $w^{i,ESG}_{t,n} = w^{ESG}_{t,n}/\sum_{n\in N^i} w^{ESG}_{t,n}$ such that $\sum_{n\in N^i} w^{ESG}_{t,n} = 1$. The residual from the projection $a_{t,n} = w_{t,n} - \sum^S_{s=1} \beta^s_t w^s_{t,n}$ is a long-short active portfolio that is orthogonal to the managed portfolios $w^s_{t,n}$. The inclusion of the equal-weighted portfolio $w^E_{t,n}$ furthermore ensures that $a_{t,n}$ is a net-zero investment portfolio, i.e. $\sum^N_{n\in N^i} a_{t,n} = 0$. The active deviation relative to the managed portfolios (as a fraction of total assets) is given by

$$\text{Active Share}_t = \frac{1}{2} \sum_{n\in N^i} |a_{t,n}| \tag{2.17}$$

Thus the projection of a fund's weights onto managed portfolios can be viewed as an extension to the 'Active Share' proposed by Cremers and Petajisto (2009). If the coefficient on the market portfolio β^{Mkt}_t is equal to 1 and the coefficients on all other managed portfolios are equal to 0, then $a_{t,n} = w_{t,n} - w^{Mkt}_{t,n}$ and the two measures of activeness coincide.

Because the weights in the zero-cost portfolio sum to 0, and all the managed portfolio weights

$w^s_{t,n}$ sum to 1 respectively, it must hold that $\sum_{s=1}^{S} \beta_t^s = 1$. The coefficients β_t^e, β_t^m and β_t^{ESG} can therefore be interpreted as the wealth shares of the individual funds $w^E_{t,n}$, $w^m_{t,n}$ and $w^{ESG}_{t,n}$ within the management company i. Figure 2.6 summarizes the ESG tilt across 13F investors. Panel (a) plots the equal- and value-weighted average ESG tilt across all 13F institutions from 2012 to 2022. The value-weighted ESG tilt β_t^{ESG} steadily grew from 7 to 18% in the past 10 years. As a robustness check, I also add 12 Fama-French industry portfolios to the projection (2.16). The aggregate ESG tilt and the corresponding total flows to the ESG portfolio are unaffected by controlling for industry exposures.

Figure 2.6: **ESG Tilts across 13F Investors**
Panel (a) plots the average equal- and value-weighted ESG tilt β_t^{ESG} across all 13F institutions. Formally, I compute $\frac{1}{I}\sum_{i=1}^{I} \beta_t^{i,ESG}$ and $\sum_{i=1}^{I} v_t^i \beta_t^{i,ESG}$ where $v_t^i = A_t^i / \sum_{i=1}^{I} A_t^i$ are AUM-weights. Panel (b) plots the fraction of Closet-ESG investors, both in terms of number of institutions and in terms of assets. Closet-ESG investors are defined as investors with an ESG-share of over 50%. The grey lines report values obtained when controlling for industry exposures in the estimation of β_t^{ESG}.

(a) Average ESG Tilt (b) Closet-ESG Investors

Using the investor-specific ESG tilts $\beta_t^{i,ESG}$ and their total assets under management A_t^i, we can compute the total ESG assets held by investor i as $A_t^{i,ESG} = A_t^i \beta_t^{i,ESG}$. Following the literature on mutual fund flows, I define the flow in the ESG portfolio of investor i as the change in ESG assets in excess of the valuation gains due to ESG returns. Formally,

$$F_{t+1}^{i,ESG} = A_t^{i,ESG} - A_t^{i,ESG}(1 + R_{t+1}^{ESG}) \tag{2.18}$$

where R_{t+1}^{ESG} is the return on the ESG portfolio. Note that empirically, this return may differ across investors because 13F institutions hold different subsets of stocks $N^i \subseteq N$. Summing across all investors yields the total flow by 13F investors in the ESG portfolio. Lastly, note that the ESG tilt β_t^{ESG} allows distinguishing 13F investors by their tilt towards sustainable stocks. I define 'Closet-ESG' investors as 13F institutions that hold over 50% of their assets in the ESG portfolio (i.e. $\beta_t^{ESG} > 0.5$). Between 2016 and 2021, the total number Closet-ESG funds grew from 8 to 133. Panel (b) plots the number of Closet-ESG funds and their total assets. The fraction of assets held by Closet-ESG funds increased over tenfold over the past 10 years.

Stock-Specific ESG Flows

If one had access to the flows and holdings of *all* investors $i = 1, ..., I$, one could decompose each investor's trades $\Delta Q^i_{t+1,n}$ into a flow-driven and an information-related component as in e.g. Greenwood and Thesmar (2011) or Lou (2012):

$$\Delta d^i_{t+1,n} = \underbrace{f^i_{t+1} Q^i_{t,n}}_{\text{Flow-Driven Demand}} + \underbrace{a^i_{t+1,n}}_{\text{Active Demand}} \tag{2.19}$$

However, the only investor group for which we have precise data on both flows f^i_{t+1} and holdings $Q^i_{t,n}$ at the fund level are mutual funds. 13F institutions only report quarterly holdings $Q^i_{t,n}$. As mentioned above, the aggregation of 13F holdings across funds within a management company makes it difficult to construct flow-driven trades for 13F investors. However, we can approximate the flow-driven trades in green stocks using the investor-specific flows in the ESG portfolio $F^{i,ESG}_{t+1}$ and the corresponding weights $w^{i,ESG}_t$.[46] For each 13F institution that is not a mutual fund, I construct $f^i_{t+1} = F^{i,ESG}_{t+1} / A^{i,ESG}_t$ as the relative flow to the ESG portfolio and $Q^i_{t,n} = (w^{i,ESG}_t A^{i,ESG}_t) / P_{t,n}$ as the corresponding stock-specific holdings within the ESG portfolio. Total flow-driven demand for each green stock is given by summing the flow-driven demand across all mutual funds and 13F investors. In order to avoid double-counting we I omit all 13F institutions that are classified as mutual funds using the corrected type codes from KY (2019).

$$\forall n \in N^G: \quad \Delta d_{t+1,n} = \sum_{i=1}^{I} Q^i_{t,n} f^i_{t+1} \tag{2.20}$$

where $N^G \subset N$ denotes the subset of green stocks. Because shares outstanding are normalized to 1, $Q^i_{t,n}$ are ownership shares and $\Delta d_{t+1,n}$ are demand shocks in percent relative to shares outstanding. Lastly, note that one could construct $\Delta d_{t+1,n}$ for alternative subsets of stocks by decomposing 13F holdings into other managed portfolios and computing the within-manager flow to the portfolio of interest.

Extracting Non-fundamental ESG Demand

Let $\mathbb{1}^X_{t,n}$ denote indicator variables equal to 1 if a trade between $t-1$ and t is an extensive trade.[47] Let $\Delta Q^X_{t,n}$ denote extensive green demand. It is the sum of all extensive trades in stock n by ESG mutual funds between $t-1$ and t:

$$\Delta Q^X_{t,n} = \sum_{i \in I^{ESG}} \mathbb{1}^X_{t,n} (Q^i_{t,n} - Q^i_{t-1,n}) \tag{2.21}$$

[46] Recall that the ESG portfolio weights are investor-specific because I normalize them within each investor's universe $w^{i,ESG}_{t,n} = w^{ESG}_{t,n} / \sum_{n \in N^i} w^{ESG}_{t,n}$.

[47] Formally

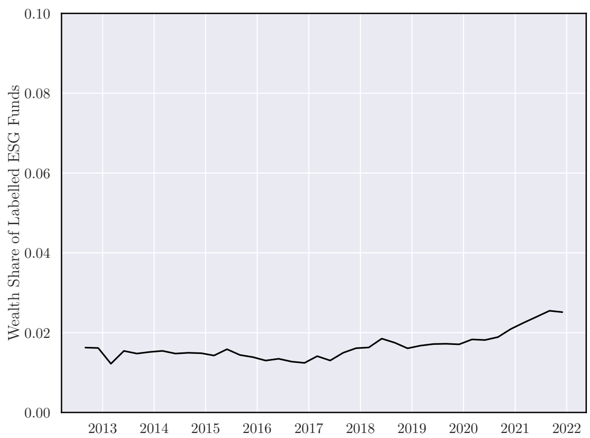
$$\mathbb{1}^X_{t,n} = \begin{cases} 0 & \text{if } Q_{t-1,n} = 0 \text{ or } Q_{t,n} = 0 \\ 1 & \text{otherwise} \end{cases}.$$

Similarly, the intensive green demand is given by $\Delta Q^{I}_{t,n} = \sum_{i \in I^{ESG}} (1 - \mathbb{1}^{X}_{t,n})(Q^{i}_{t,n} - Q^{i}_{t-1,n})$. As described above, the goal is to extract the mandate-driven (i.e. non-fundamental) component of $\Delta Q^{X}_{t,n}$. Under the assumption that all fundamental information contained in extensive trades is also present in intensive trades, we can extract the mandate-driven component from extensive green demand $Q^{X}_{t,n}$. To this end, I cross-sectionally orthogonalize extensive green demand with respect to intensive green demand,

$$\forall t \quad : \Delta Q^{X}_{t,n} = \beta_t \Delta Q^{I}_{t,n} + Controls + \Delta Q^{\perp}_{t,n} \tag{2.22}$$

where $Controls$ includes changes in book equity, total assets and profitability between $t-1$ and t. One could argue, that intensive trades by ESG funds do not capture all of the fundamental information in extensive trades. I therefore also orthogonalize with respect to total intensive trades (i.e. trades summed across all investors and not just ESG funds). This should eliminate any variation in ESG funds' extensive trades that is driven by fundamental information. The residual from the regressions, $\Delta Q^{\perp}_{t,n}$, is the component of ESG funds' purchases that is exclusively driven by exogenous portfolio constraints or mandates. That is, it is a proxy for non-fundamental ESG demand. However, because $\Delta Q^{\perp}_{t,n}$ is constructed using only a subset of all ESG investors, regressions of returns onto $\Delta Q^{\perp}_{t,n}$ do not capture structural parameters. In order to approximate total mandate-driven ESG demand, let S^{ESG}_{t} denote the wealth share of labelled ESG mutual funds relative to total ESG assets A^{ESG}_{t}. Figure 2.7 plots S^{ESG}_{t} over time.

Figure 2.7: **Wealth Share of Labelled ESG Mutual Funds**
The figure reports the wealth share of labelled ESG mutual funds as a fraction of total ESG assets A^{ESG}_{t}. Total ESG assets are constructed as the sum of individual investors' ESG holdings $A^{i,ESG}_{t}$ from Section 3.3.



The figure suggests that demand from labelled ESG mutual funds is roughly 2% of total ESG demand. Total mandate-driven ESG demand can be approximated as $\frac{1}{S^{ESG}_{t}} \Delta Q^{\perp}_{t,n}$. The coefficient obtained from regressions of returns onto total mandate-driven ESG demand should then reveal the structural parameter. Lastly, note that for simplicity of exposition, the estimation is split into the construction of $\Delta Q^{\perp}_{t,n}$ and regressions of returns onto $\Delta Q^{\perp}_{t,n}$. This is equivalent to simple regressions of returns onto the raw $\Delta Q^X_{t,n}$, controlling for intensive trades $\Delta Q^I_{t,n}$.

C Identification

Dividend Reinvestments

Do mutual funds reinvest total dividend payout in their existing portfolio? I assess the extent to which mutual funds invest a stock's dividend payout in all other stocks within their portfolios. Let $\Delta q^i_{t,n} = Q^i_{t,n}/Q^i_{t-1,n} - 1$ denote the percentage change in shares held between two quarters. If mutual funds reinvest dividend payouts across their entire portfolio, then $\Delta q^i_{t,n}$ should be significantly related to the dividend flow from all *other* stocks $df^i_{t,-n}$. I test this in a pooled regression given by

$$\Delta q^i_{t,n} = \theta df^i_{t,-n} + Controls + \epsilon^i_{t,n} \tag{2.23}$$

where $Controls$ includes a constant, time fixed effects, total fund flows f^i_t and log returns $\Delta p_{t,n}$. Table 2.13 reports the coefficient estimates across different specifications. The dividend-

Table 2.13: **Dividend Reinvestments**
The table reports the estimated coefficients from the pooled regression of trades $\Delta q^i_{t,n}$ onto dividend flows from other stocks $df^i_{t,-n}$. Standard errors robust to heteroskedasticity and autocorrrelation are reported in parentheses. Significance at the 90, 95 and 99% confidence levels is indicated by *,**,*** respectively.

	Quarterly trades $\Delta q^i_{t,n}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Dividend Flow $df^i_{t,-n}$	**4.88*** **	**5.83*** **	**2.52*** **	**2.44*** **	**1.21** **	**1.91*** **
	(0.807)	(0.836)	(0.545)	(0.408)	(0.502)	(0.554)
Total Flow f^i_t	-	-	0.61	0.67	0.27	0.21
	-	-	(0.046)	(0.035)	(0.027)	(0.025
Stock Return $\Delta p_{t,n}$	-	-	-0.02	-0.03	-0.04	-0.04
	-	-	(0.003)	(0.003)	(0.012)	(0.018)
Div. per Share $D_{t,n}$	-	-	-	-0.01	0.02	-
	-	-	-	(0.001)	(0.008)	-
Quarter FE	No	Yes	Yes	Yes	Yes	Yes
Large Dividend Flow ($df^i_{t,-n} > 1\%$)	No	No	No	No	Yes	Yes
Exclude Dividend Stocks	No	No	No	No	No	Yes

scaling coefficient θ is significantly positive across all specifications. Thus, on average mutual funds reinvest their dividend payouts across all other stocks in their portfolios. Specification (5) estimates the impact of large dividend flows that exceed 1% of the fund's total assets. The scaling coefficient is close to 1, suggesting that when funds receive large total dividend flows, they proportionately scale up all their existing positions. Specification (6) excludes dividend-paying stocks. While this greatly reduces the number of observations, the dividend scaling coefficient remains highly statistically significant. Lastly, note the effect of total relative flows f_t^i on quarterly trades is highly statistically significant across all specifications. The estimated coefficient is comparable in magnitude to Lou (2012). Mutual funds scale their existing portfolio holdings in response to both, total in- and outflows, and total payouts from dividends.

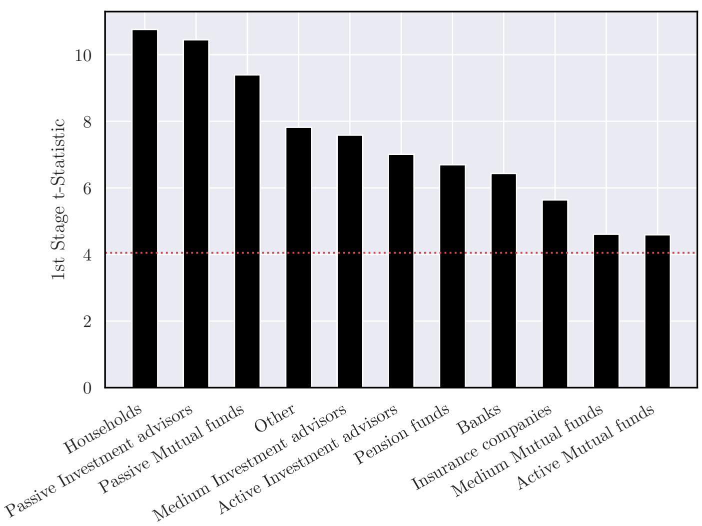
Instrument Relevance

The significant relationship between flow-driven purchases and contemporaneous returns (i.e. the relevance of the instrument) has been shown at least since Lou (2012). Schmickler and Tremacoldi-Rossi (2022) furthermore show that the trades induced by dividend flows are significantly related to contemporaneous stock returns. This paper uses their instrument to identify exogenous price shocks in the first stage. Recall, that the first stage regression is given by

$$\Delta p_{t,n}^i = \theta^i DIT_{t,n}^{-i} + \epsilon_{t,n}^i$$

where $DIT_{t,n}^{-i}$ is the dividend-based instrument and $\epsilon_{t,m}^i$ includes the control variables log book equity, profitability, investment, and market beta. I estimate the first stage in a pooled panel regression for the subset of stocks held by Thomson Reuters' institutional type codes. Figure 3.3 reports the t-statistic of θ^i for each institutional type. The t-statistic exceed the critical weak-instrument threshold of of 4.05 (see Stock and Watson, 2005) for all institutional types.

Figure 2.8: **Weak Instrument Test**
The figure reports the t-statistic on the coefficient θ^i for each institutional type. The red dotted line indicates the weak instrument threshold of 4.05 (see Stock and Watson, 2005).



Identification using Benchmarking Intensity

The central object of the relationship between flow-driven demand and realized returns is investors' elasticity of demand. The parameter estimates critically depend on a valid instrument, i.e. *exogenous* variation in prices that is orthogonal to the investor's own latent demand shocks. In this section I explore the stability of the estimates to an entirely different instrument. In particular, I use the benchmarking intensity by Pavlova and Sikorskaya (2022) as exogenous supply shocks to identify investors' demand elasticities. I kindly thank Anna Pavlova and Taisiya Sikorskaya for generously sharing their data with me.

Benchmarking Intensity. If the aggregate demand for equities is downward-sloping, changes in the supply of stocks can have significant price effects. These supply shifts can in turn be used to identify investors' demand elasticities. If benchmarked (or passive) investors increase their holdings in a given stock, they are effectively reducing its supply. Pavlova and Sikorskaya (2022) construct each stock's benchmarking intensity as the AUM-weighted sum across index weights:

$$BMI_{t,n} = \frac{\sum_x A_t^x w_{t,n}^x}{P_{t,n}} \tag{2.24}$$

where $w_{t,n}^x$ is the portfolio weight of stock n in index x at time t and A_t^x is the total amount of exchange traded funds' and mutual funds' assets benchmarked to index x. Changes in a stock's benchmarking intensity between two quarters, $\Delta BMI_{t,n}$, represent a change in the stock's effective supply and may be used as an exogenous shock to identify elasticities. However, changes in benchmarking intensity may be driven by fundamental news (or price increases themselves) that cause index additions and deletions. Pavlova and Sikorskaya (2022) address the potential endogeneity by only using changes in benchmarking intensity across the Russell 1000/2000 cutoff. However, the amount of observations around the Russell cutoff are not sufficient to estimate the structural model in Section 3.4. Despite the potential endogeneity concerns, I therefore use the raw $\Delta BMI_{t,n}$ across all stocks and quarters to identify investor-specific demand elasticities. While this instrument is imperfect, it serves as a useful robustness check that similar elasticity estimates can be obtained using a completely different instrument from a separate study.

Table 2.15 summarizes the impact of ESG demand shocks identified from changes in benchmarking intensities. The estimates are close to the estimates identified from the flow-based instrument (see Table 2.3). The average price impact of the demand for green stocks is 1.17% respectively. Recall, that the multiplier identified from the flow-based instrument is 1.11%. Thus the multiplier identified via benchmarking intensity is slightly higher but lie in the same ballpark.

Table 2.14: **Multiplier Matrix identified from Benchmarking Intensity**
The table summarizes the price impact $\mathcal{M}$ of green demand onto green g and non-green stocks b. (i.e. diagonal and off-diagonal elements of the multiplier matrix $\mathcal{M}$) identified from from changes in Benchmarking Intensity $\Delta BMI_{t,n}$ (see Pavlova and Sikorskaya (2022).

		Cross-Multipliers ($\times 10^4$)	
	$\mathcal{M}_g$	$\mathcal{M}_{gg}$	$\mathcal{M}_{bg}$
Mean	1.17	-0.82	1.50
Std.	(1.06)	(29.48)	(33.99)
10th Pctl.	0.35	-8.46	-6.27
Median	0.83	0.00	0.18
90th Pctl.	2.5	6.61	10.87
Fraction Positive Spillovers		50%	61%

Identification using Flow Shocks

Following van der Beck (2022), I construct an alternative instrument for each investor by aggregating the surprise flow-induced trades by all other mutual funds,

$$f_{t,n}^{-i} = \sum_{j \neq i}^{I} f_{t+1}^{j,\perp} Q_{t,n}^{j} \tag{2.25}$$

where $f_{t+1}^{j,\perp}$ is the flow into fund j between t and $t+1$ orthogonalized for fund characteristics, holdings and returns. In particular, I obtain the surprise flow $f_t^{j,\perp}$ as the residual in cross-sectional regressions of fund flows f_t^j onto fund characteristics. The characteristics are portfolio weighted greenness, value, size, momentum, profitability, investment and idiosyncratic volatility as well as the funds' own contemporaneous returns. Thus the flow shocks are orthogonal to the fund's portfolio tilts. This addresses the endogeneity concern that flows are driven by fundamental news regarding the fund's underlying assets. See van der Beck (2022) for details.

D Details on Estimation and Variable Construction

Pooling by Institutional Types

In order to pool investors into groups, I start by computing each investor's active share as of quarter t as

$$\text{Active Share}_t^i = \frac{1}{2} \sum_n |w_{t,n}^i - w_{t,n}^M| \tag{2.26}$$

Table 2.15: **Multiplier Matrix identified from Flow Shocks**
The table summarizes the price impact $\mathcal{M}$ of green demand onto green g and non-green stocks b. (i.e. diagonal and off-diagonal elements of the multiplier matrix $\mathcal{M}$) identified from flow shocks (see van der Beck (2022).

		Cross-Multipliers ($\times 10^4$)	
	$\mathcal{M}_g$	$\mathcal{M}_{gg}$	$\mathcal{M}_{bg}$
Mean	2.78	-16.56	-9.81
Std.	(0.72)	(62.99)	(40.27)
10th Pctl.	1.81	-41.63	-24.23
Median	2.82	-3.68	-1.31
90th Pctl.	3.67	0.00	0.56
Fraction Positive Spillovers		10%	21%

which measures a fund's deviation from holding a passive market portfolio. I define index funds as 13F institutions with Active Share$_t^i < 0.01$. These are all investors who tilt less than 1% of their portfolio away from passive market weights. For the remaining investors, I use Thomson's institutional type code labels, which split investors into banks, pension funds, investment advisors, insurance companies, mutual funds and other. I divide the largest investor groups, investment advisors and mutual funds, into activeness terciles based on their Active Share$_t^i$. The resulting groups are labelled *rigid*, *medium* and *elastic*. I estimate group-specific demand curves using the two-step procedure in (2.12) by pooling the observations of all institutions within a group.

Estimated Coefficients by Investor Type

I estimate elasticities over the panel of quarterly holdings from 2010 to 2020 including time fixed effects. Table 2.16 reports the estimated coefficients for all investors. The first row reports the estimates for a pooled regression across all investors. The pooled elasticity is 1.05, which implies that on average institutions sell 1.05% of their holdings in a stock when the price increases by 1%.[48] The remaining rows report the elasticities obtained in a pooled estimation across institutional types. There is great heterogeneity in the estimated elasticities ζ^i across types. ζ^i is the lowest for insurance companies and large passive investment advisors investors such as Blackrock, Fidelity and Vanguard. Active mutual funds are the most elastic investors with an elasticity of 3.2. The second column reports the elasticities estimated from the cross-section of quarterly holdings $Q_{t,n}^i$ instead of trades $\Delta q_{t,n}^i$. The estimates are considerably smaller for the majority of investors.

[48]Note. that ζ^i only approximates the true elasticity. The next section provides a thorough description on how to structurally construct exact elasticities.

Table 2.16: **Demand Curves by Investor Type**

The table reports the estimated demand curves for different groups of investors. The trades $\Delta q^i_{t,n}$ are pooled over stocks, quarters and institutional types, such that one demand curve is estimated per investor group. Formally, the estimation equation is given by

$$\forall j = 1,...,J \quad : \Delta q^j_{t,n} = -\zeta^j \Delta \hat{p}_{t,n} + \epsilon_{t,n}.$$

where $\Delta \hat{p}_{t,n}$ is the fitted return from dividend flow demand shocks and $\epsilon_{t,n}$ includes the control variables log book equity, profitability, investment and market beta. Institutional types (split by active share) are denoted by $j = 1,...J$ and include mutual funds, investment advisors, households, pension funds, insurance companies, and other 13F institutions. Standard errors (in parentheses) are robust to heteroskedasticity and autocorrelation.

	ζ^i Identified from Trades $\Delta q_{t,n}$ van der Beck (2022)	ζ^i Identified from portfolio holdings $Q_{t,n}$ Koijen and Yogo (2019)
Pooled All		
	1.054 (0.033)	0.282 (0.001)
Pooled by Type		
Mutual Funds		
High Active Share	3.198 (0.305)	0.744 (0.004)
Medium Active Share	2.660 (0.298)	0.477 (0.004)
Low Active Share	1.296 (0.092)	-0.142 (0.003)
Investment advisors		
High Active Share	0.924 (0.120)	0.795 (0.002)
Medium Active Share	0.250 (0.103)	0.624 (0.001)
Low Active Share	0.424 (0.046)	0.521 (0.001)
Banks	1.292 (0.118)	0.238 (0.002)
Pension funds	0.838 (0.081)	0.322 (0.002)
Insurance companies	0.387 (0.168)	0.321 (0.003)
Other 13F Instiutions	0.039 (0.226)	-0.415 (0.003)
Households	0.724 (0.244)	0.530 (0.009)

Incorporation in Logit Framework and Asset Substitution

Motivated by the fact that portfolio weights are log-normally distributed in the data, KY (2019) propose (and microfound) a logit framework for the demand of investor i:

$$\log\delta^i_{t,n} = (1-\zeta^i)\log P_{t,n} + \varepsilon^i_{t,n} \tag{2.27}$$

where $\delta^i_{t,n} = w^i_{t,n}/w^i_{t,0}$ is the portfolio weight relative to the weight in an outside asset $w^i_{t,0}$ and $\varepsilon^i_{t,n}$ includes a constant, observable characteristics and a residual. The portfolio constraint that $\sum_{n\in N^i} w^i_{t,n} = 1 - w^i_{t,0}$ implies that

$$w_{t,n} = \frac{\delta^i_{t,n}}{1+\sum_{m=1}^N \delta^i_{t,m}}. \tag{2.28}$$

The logit framework ensures that portfolio holdings add up to total assets and that holdings cannot be negative (as observed in the 13F filings). Note that we can rewrite $\log\delta^i_{t,n} = \log Q^i_{t,n} + \log P^i_{t,n} - \log w^i_{t,0}A^i_t$. Rearranging and taking first differences yields[49]

$$\Delta q^i_{t,n} = -\zeta^i_t \Delta p_{t,n} + \epsilon^i_{t,n} \tag{2.29}$$

where $\epsilon^i_{t,n} = \Delta\log(w^i_{t,0}A^i_t) + \Delta\varepsilon^i_{t,n}$. The first-difference estimator has the distinct advantage of eliminating any time-invariant drivers of cross-sectional portfolio holdings that are correlated to prices. In a simple simulation, van der Beck (2022) shows that (2.29) successfully eliminates the ommitted variable bias from unobservable portfolio tilts that are slow-moving and correlated to the cross-section of prices.

A first-order approximation of investor's demand elasticity is given by the scalar regression coefficient $-\frac{\Delta q^i_{t,n}}{\Delta p_{t,n}} = \zeta^i$. However, this measure of elasticity does not ensure that the investor's portfolio weights add up to 1, (or alternatively: that her assets A^i_t remain unchanged). In order to ensure that the budget constraint holds we need to plug the estimated coefficient into (2.28). To this end, note that $\log w^i_{t,n} = \log Q^i_{t,n} + \log P^i_{t,n} - \log A^i_t$. Differentiating and rearranging yields the following elasticity

$$-\frac{\partial\log Q^i_{t,n}}{\partial\log P_{t,n}} = \zeta^i + \underbrace{w^i_{t,n}(1-\zeta^i)}_{\text{Portfolio Constraint}}. \tag{2.30}$$

The elasticity is given by ζ^i plus a correction term, which ensures that portfolio weights add up to 1.[50] Precisely because of the portfolio constraint, price changes have spillover effects to

[49] KY (2019) actually propose re-estimating 2.27 over the cross-section of portfolio weights every quarter t resulting in time-varying coefficients ζ^i_t. Empirically, however, the coefficients remain very stable in the time-series. Thus the correction term for time-varying coefficients in the first-difference estimator is small and can be ignored. In fact, estimating constant demand coefficients ζ^i in a panel regression including time-fixed effects leads to essentially the same demand curves (see van der Beck and Jaunin (2021) and Koijen et al. (2022))

[50] The correction term is negative, if the investor is very elastic $\zeta^i > 1$. In this case the dollar holdings (not the

other stocks. Cross-elasticities are given by

$$-\frac{\partial \log Q^i_{t,n}}{\partial \log P_{t,m}} = w^i_{m,t}(1-\zeta^i).^{51} \tag{2.31}$$

We can stack the elasticities into an elasticity matrix $-\frac{\partial \log Q^i_t}{\partial \log P'_t} \in \mathbb{R}^{N\times N}$ given by

$$-\frac{\partial \log Q^i_t}{\partial \log P'_t} = \zeta^i I + (1-\zeta^i)1w'_t, \tag{2.32}$$

where I is the identity matrix and 1 is a vector of ones. Thus the logit framework allows transforming the simple scalar regression coefficient ζ^i into a demand-elasticity matrix that accounts for spillover effects across the entire cross-section of holdings.

Details on the Flow Simulation

Let $\Delta P^{\text{ESG-Flow}}_{t+1}$ denote vector of price pressures (expressed in dollars) resulting from \$X flow from the market portfolio towards the ESG portfolio. Equation (2.7) implies that

$$\Delta P^{\text{ESG-Flow}}_{t+1} = \mathcal{M}_t(w^{ESG}_t - w^{MF}_t) * \$X$$

Note, that (2.7) is expressed in percentage terms (i.e. the return $\Delta p_{t+1,n}$ resulting from a demand shock in percent of shares outstanding). It can also be expressed in terms of dollar terms by multiplying by prices $P_{t,n}$ (which are equal to market equities due to the normalization). The price pressure in percentage terms for each stock n is given by

$$\Delta p^{\text{ESG-Flow}}_{t+1,n} = \frac{\Delta P^{\text{ESG-Flow}}_{t+1,n}}{P_{t,n}}$$

Note, that true (empirically observed) realized returns of the ESG portfolio are given by $R^{ESG}_{t+1} = \sum_n \tau_{t,n} r_{t+1,n} = \tau'_t r_{t+1}$. The structurally implied price pressure from \$X ESG flows is given by

$$Pressure^{ESG}_{t+1} = \sum_{n=1}^{N} \tau_{t,n} p^{\text{ESG-Flow}}_{t+1,n}$$

A first order approximation of the counterfactual ESG returns in the absence of flow-driven price pressure is therefore given by

$$\tilde{R}^{ESG}_{t+1} = R^{ESG}_{t+1} - Pressure^{ESG}_{t+1}$$

number of shares held!) in stock n is decreasing in price of n and we have to make a downward adjustment to the elasticity to satisfy the portfolio constraint.

[51] Again, if $\zeta^i > 1$ an increase in the price of stock m reduces the dollar holdings in stock m and the freed up cash is invested in all other stocks, causing spillover effects proportional to the size of the shock to m given by $w^i_{m,t}$.

. Table 2.17 reports the ESG portfolio's counterfactually observed alpha in the absence of price pressure arising from simulated sustainable flows of $10 and $25 billion quarterly.

Table 2.17: **ESG Alpha without Flow-driven Price Pressure**
The table reports annualized long-short returns and alphas of the ESG portfolio from 2016 to 2021. On the left I report the empirically observed alphas. On the right, I report the counterfactually observed alphas without the price pressure from simulated sustainable flows. The long-short ESG portfolio τ_t is the zero-investment portfolio that goes long the ESG portfolio w_t^{ESG} and short the aggregate mutual fund portfolio w_t^{MF}. Alphas are computed with respect to the CAPM, the CAPM plus the Green Factor in PST (2022), and Carhart 4-factor model plus the Green Factor. The standard errors are robust to heteroskedasticity and autocorrelation.

	Return	α (CAPM)	α (CAPM + Green)	α (CH4 + Green)
True Returns: Empirically Observed				
Return (%)	2.01	2.40	1.87	1.51
t-statistic	2.91	3.47	2.55	2.01
Counterfactual Returns: In Absence of Flows of Simulated Flow $25B				
Return (%)	0.66	1.05	0.5	0.12
t-statistic	0.86	1.36	0.6	0.14

The price pressure from quarterly ESG flows of $10 billion is already sufficient to account for almost all of the outperformance of ESG funds. Furthermore, in the absence of $25 billion quarterly ESG flows, the counterfactual returns and alphas of the ESG-taste portfolio are all negative. In contrast, the realized (i.e. truly observed) returns and alphas are all significantly positive. These results emphasize the sizeable gap between realized and expected returns from ESG investing that is driven by flows to sustainable funds. This suggests that without continued flow to sustainable funds, ESG investing may have negative alpha. In other words, it is the price pressure from ESG flows that made 'doing well by doing good'-investing possible.

Variable Construction

- Book Equity: Book equity is constructed following Fama and French. It is the book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit, minus the book value of preferred stock.
- Market beta: Stocks' market betas are estimated in a regression of monthly returns over the one-month T-bill. We use 60-month rolling windows and require a minimum of 24 months of returns.

- Profitability: I use the Fama and French definition, i.e. revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses divided by book equity for the last fiscal year end in $t-1$.
- Investment: Investment is the annual growth rate of assets computed as a log difference.
- Idiosyncratic volatility: Idiosyncratic volatility is computed as the monthly time series standard deviation of residual returns. Residual returns are obtained from regressing daily returns onto daily realization of the market, size and value factor.
- Turnover: Turnover is the total share volume in a given month (as reported on CRSP) divided by the total shares outstanding.
- Momentum: Momentum is computed as the total return over the past 11-months, excluding the most recent month.
- Sin Stocks: Tobacco and alcohol stocks are defined as belonging to the Fama and French SIC classification groups 4 and 5 respectively. Gaming stocks are identified using NAICS codes 7132, 71312, 713210, 71329, 713290, 72112, and 721120.
- MSCI Controversial Stocks: Stocks in the biotech, firearms, oil, military and cement industry are identified using SIC codes 2833–2836, 1300, 1310–1339, 1370–1382, 1389, 2900–2912, 2990–2999, 3240–3241, 3760–3769, 3795, 3480–3489 and NAICS codes 336992, 332992–332994.
- Co2 Emissions: Total Co2 Emissions are scaled by revenue. As the fraction of US stocks with available Co2 emissions is relatively small, I compute an additional proxy for Co2 emissions which fills missing observations with the median emissions within Fama and French 48 industries.

Chapter 3

The Equity Market Implications of the Retail Investment Boom

3.1 Introduction

The emergence of commission free brokerage has changed the investment landscape for small retail investors. The online trading platform "Robinhood Markets Inc." was the first brokerage to offer costless trading to households - a large group of potentially less sophisticated investors. During the onset of the Covid-19 pandemic, its users (henceforth Robinhood traders) have received considerable attention by the media.[1] A reoccurring narrative has been that retail trading activity may be the driver of many sudden surges in equity valuations observed during the first half of 2020. A popular example is the car rental company Hertz. Shortly after it filed for bankruptcy on May 22nd, 140,000 Robinhood traders added Hertz to their portfolios. Hertz's share price simultaneously jumped by over 1,000%. However, it was the beginning of 2021 that put retail traders' price impact at the center of attention. On "r/wallstreetbets", which is a subforum on the message board Reddit, retail traders coordinated to collectively purchase shares and call options of several companies they deemed undervalued. Within the second half of January 2021, the share price of the video game retailer GameStop rose from 20$ to 486$. There exists widespread belief that retail traders can affect the price of small companies such as Hertz or GameStop.[2] However, given their small assets under management, it seems unlikely that retail investors can affect the aggregate US stock market with a market capitalization of over $40 trillion, or the share price of large individual companies such as General Electric, Ford or General Motors. Furthermore, it is unclear why and over which

[1] The Wall Street Journal, July 25, 2020: *"Everyone's a Day Trader Now"*; Financial Times, August 17, 2020: *"Retail trading app Robinhood's value tops $11bn on new fundraising"*; Bloomberg News, July 13, 2020: *"Ten Thousand Day Traders an Hour Are Buying Tesla Shares"*; Business Insider, July 14, 2020: *"Nearly 40,000 Robinhood day traders added Tesla shares in 4 hours Monday as the stock whipsawed"*

[2] Bloomberg News, June 12, 2020: *"Day Traders Might Have Fun Saving Hertz From Bankruptcy"*; Financial Times, January 22, 2021: *"GameStop shares leap as day-trading 'mob' tussles with short seller"*

horizon non-fundamental retail demand shocks impact prices.

Motivated by these events, we seek to quantify the impact of retail traders' demand on the US equity market within a structural model of asset demand. In our analysis, we specifically consider the demand of Robinhood traders. We focus on the onset of the Covid-19 crisis from January to July 2020. The first two quarters of 2020 (henceforth Q1 and Q2) are well-suited to assess the impact of retail investors' demand. In early 2020, the pandemic came as an exogenous shock which altered the perspectives of many firms and induced financial institutions to rebalance their portfolios. Simultaneously, retail trading activity - presumably fueled by commission free brokerage, stimulus checks and a lack of consumption opportunities - increased sharply. This paper investigates the dual role played by Robinhood traders in accommodating market-wide supply shocks during Q1 and amplifying the extraordinary recovery in Q2. We motivate our analysis in reduced form by showing a positive relationship between Robinhood demand and stock returns during Q2. We identify Robinhood demand using a novel dataset of account holdings from "Robinhood Markets Inc." provided by Robintrack. We find that the abnormal returns of a portfolio that longs high demand stocks and shorts low demand stocks are significantly greater than zero from April to June. Furthermore, we document a significantly positive cross-sectional relationship between stock returns and Robinhood demand, controlling for a range of stock-specific characteristics. Based on these findings we investigate the following question: To what extent did Robinhood demand affect prices during the stock market crash in Q1 and the subsequent recovery in Q2?

As reduced-form regressions cannot answer this question, we estimate a demand system along the lines of Koijen and Yogo (2019) (henceforth KY) to quantify the impact of the retail investment boom. Their "Demand System Approach to Asset Pricing" accounts for the endogeneity of demand and prices, and is therefore able to elicit the *causal* effect of Robinhood demand on prices. We approximate the Robinhood portfolio using data from Robintrack and institutional portfolios using 13F filings. The residual shares not held by 13F institutions or Robinhood traders make up the household portfolio as in KY (2019). Since institutional holdings are only available quarterly, we restrict our analysis to this frequency. We stress that an exact quantification of Robinhood traders' price impact is impaired by the fact that their *true* holdings remain unobserved. However, we provide a reasonable approximation of the Robinhood portfolio using the number of Robinhood accounts invested in each stock, which is publicly available on Robintrack. We emphasize that this approximation is the best one can do given the currently available data on retail and institutional investors' equity holdings.[3] We then estimate characteristics-based demand curves for institutional investors, households and Robinhood traders, taking into account the endogeneity of prices and latent demand. In contrast to KY (2019), our focus lies on running counterfactual experiments. More specifically, we want to assess the counterfactual equilibrium observed in the absence of Robinhood traders. KY (2019) re-estimate demand quarterly and hence obtain time-varying demand curves. In a deviation from their methodology, we estimate demand over the panel

[3]We furthermore provide simulations which suggest that our qualitative results are not driven by potential misapproximations.

of observed portfolio holdings and obtain time-invariant demand curves. This substantially increases the number of observations for each investor and therefore raises the statistical significance of the estimated demand curves. It furthermore facilitates the incorporation of Robinhood traders' demand curves in the institutional demand system. The responsiveness of investors' portfolio allocation to the price is given by the estimated demand coefficient on the characteristic *market equity*. This coefficient drives an investor's demand elasticity with respect to the price. We find an extremely low price elasticity for institutional investors, who hold roughly 65% of the US equity market. In fact, 39% of all invested wealth is held by passive institutional investors that are fully inelastic. This inelasticity of equity markets, which is potentially driven by investment mandates and the rise of passive investing, is in line with recent findings by **?**. We find that it is the inelastic demand response of institutional investors that allows Robinhood traders' demand shocks to have sizeable price effects.

Based on the estimated demand curves and assets under management (AUM) disclosed in Robinhood's S-1 filing, we impose market clearing to derive an equilibrium price for each stock, which is a function of exogenous stock-specific fundamentals as well as the estimated demand coefficients and AUM of institutions, households, and Robinhood traders. We use the market clearing condition in three counterfactual experiments. First, we decompose the cross-sectional variation in log returns during Q1 and Q2 into supply- and demand-driven components. We find that in Q1 much of the variation in stock returns was driven by *real* effects, i.e. changes in fundamentals. During the recovery in Q2 however, the majority of return variation was driven by demand effects, i.e. flows between investors. Despite their negligible market share, we show that Robinhood traders accounted for 10% of the cross-sectional variation in stock returns.

Second, we estimate a counterfactual US equity market in the absence of Robinhood traders in 2020. To this end, we redistribute Robinhood traders' assets across all other investors during Q1 and Q2. The resulting counterfactual market prices approximate what equity valuations would have been without the retail investment boom. Our findings suggest that Robinhood traders were providing considerable liquidity to the equity market during the crash and amplified the recovery in Q2. Decomposing the crash and recovery into size quintiles, we find that Robinhood traders' price impact is concentrated towards smaller stocks. In July 2020, Robinhood demand accounted for 25% of the aggregate market capitalization of the smallest size quintile and over 40% of the market capitalization of the smallest size decile. The price impact decreases monotonically moving to larger percentiles. Furthermore, the aggregate market capitalization of both energy and industrial stocks would have been almost 1% lower without the presence of Robinhood traders. We compute standard errors around the counterfactual equilibrium via simulating alternative demand curves and find that both the estimated elasticities as well as Robinhood traders' equilibrium price impact are highly statistically significant.

Third, we examine the effects of Robinhood traders on the time-series volatility of individual stocks. To this end, we extend the demand system to jointly match prices and Robinhood holdings at a daily frequency. We find, that except for a few idiosyncratic herding events, Robinhood trading activity lowers the volatility for the majority of all stocks. In particular, 71% of the stocks not directly traded on Robinhood have a lower volatility due to the presence of Robinhood traders.

Fourth, we assess the return implications of the retail investment boom for individual stocks. We show that Robinhood traders' price impact can be approximated as a product of their ownership share and an inelasticity multiplier. A stock's inelasticity multiplier is determined by the ownership-weighted demand elasticity of its investor base. Robinhood traders hold a relatively small fraction of the total shares outstanding for stocks within the S&P500. We show that they nevertheless provided considerable liquidity for some large individual stocks in Q1 and strongly amplified their subsequent recovery in Q2. Robinhood traders' outsized price impact is driven by the fact that these stocks have high multipliers, as they are primarily held by large passive investors with strongly inelastic demand curves. These findings shed light on the recent events in January 2021. While GameStop was subject to substantial short interest by hedge funds, its long investor base was primarily inelastic. We find that, as of July 2020, the inelasticity multiplier was 5.5. This implies that, ceteris paribus, buying 10% of GameStop's shares outstanding results in a 55% price increase.

Related Literature This paper is located within the recent and growing literature on the inelastic nature of equity markets. This literature has its roots in some early attempts to estimate the slope of aggregate equity demand (Shleifer, 1986; Wurgler and Zhuravskaya, 2002). In their "Demand System Approach to Asset Pricing", KY (2019) propose a practical framework to estimate investor-specific demand functions and, imposing market clearing, derive equilibrium prices consistent with these demand functions. In a follow-up paper, Koijen et al. (2022) link global valuation ratios to the estimated demand coefficients of different investor groups. While we closely follow the methodology of KY (2019), we adopt the panel estimation from Koijen et al. (2022) in order to obtain time-invariant demand curves which increases statistical significance and facilitates the incorporation of Robinhood traders' demand. In a closely related paper, Gabaix and Koijen (2021) estimate the macro-elasticity of demand for the aggregate stock market and find that investing $1 in the market portfolio increases the aggregate stock market capitalization by $5. We, on the other hand, are concerned with the micro-elasticity of demand and Robinhood traders' impact on the cross-section of stock returns.

More generally, this paper relates to the large literature analyzing the trading behavior of retail investors. Barber et al. (2008) find that stocks that retail investors are buying (selling) exhibit positive (negative) abnormal returns in the current and following weeks. These findings are confirmed by Boehmer et al. (2020) who find that retail order flow predicts the cross-section of stock returns over the following three weeks. In line with these findings, we show that the stocks bought by Robinhood traders display positive abnormal returns in Q2. Barrot et al. (2016) show that the retail sector provides liquidity to the equity market in times of stress. Likewise, we find that Robinhood traders acted as liquidity providers during the stock market crash in March 2020.

Our paper furthermore contributes to the recent literature analyzing the trading behavior of Robinhood users. Barber et al. (2020) find that the information displayed on the Robinhood trading platform incites investor herding to a level that can impact prices.[4] Glossner et al. (2020) and Welch (2021) find suggestive evidence that Robinhood traders provided liquidity during Q1.[5] We quantify their findings and furthermore show that Robinhood also contributed positively to the surge in equity prices in Q2. Lastly, Eaton et al. (2021) use intraday data to show that Robinhood platform outages are associated with higher market liquidity among high attention stocks favored by Robinhood traders. They argue that zero-commission investors increase market makers' inventory risk. This hypothesis is distantly related to ours, as inelastic market participants can be viewed as a source of inventory risk for market makers.

The remainder of this paper is structured as follows. Section 3.2 describes the stock market and holdings data. Section 3.3 motivates our investigation via a reduced-form analysis of the relationship between Robinhood demand and stock returns. Section 3.4 describes the demand system and the estimation procedure. Section 3.5 presents our empirical results on the impact of Robinhood traders on the equity market. Section 3.6 concludes.

3.2 Data

A Stock Characteristics

While this study focuses on the impact of Robinhood demand on the US equity market during the first half of 2020, the estimation of demand curves is based on data reaching back to 2005. The accounting and financial data used in this paper are obtained from Refinitiv. Our initial sample of stocks is comprised of all ordinary common shares traded on the New York Stock Exchange (NYSE), the NYSE American, and NASDAQ between January 2005 and June 2020.[6] We follow KY (2019) and use log market equity, log book equity, profitability, investment, dividends to book equity, and market beta as stock-specific characteristics in the demand estimation. Profitability is the ratio of operating profits to book equity. Investment is the annual growth rate of assets computed as a log difference. Dividends to book equity is the annual split-adjusted dividend divided by book equity. Stocks' market betas are estimated in a regression of monthly returns over the one-month T-bill.[7] Finally, profitability, investment, and market beta are winsorized at the 2.5th and 97.5th percentile in the cross-section, while dividends to book equity is winsorized at the 97.5th percentile. We construct an outside asset from the set of firms which are either not headquartered in the US, are real estate investment

[4] In a complementary paper, Ozik et al. (2020) provide evidence that attention-grabbing stocks were more heavily traded on Robinhood during the Covid-19 crisis.

[5] Similarly, Pagano et al. (2020) show that Robinhood traders engaged in contrarian trading during the crisis.

[6] We ignore stocks with missing shares outstanding and prices.

[7] We use 60-month rolling windows and require a minimum of 24 months of returns.

trusts (i.e, GICS classification 601010) or have missing characteristics. Excluding the outside asset, our final investment universe contains 2,078 stocks in Q1 and 2,117 in Q2. When we refer to time spans, we use Q1 and Q2 to denote the time span from January 1st to March 31st, 2020 and April 1st to June 30th, 2020 respectively. When we refer to dates, we use Q1 and Q2 to denote March 31st and June 30th. These dates are chosen in line with the reporting dates at which 13F institutions have to file their holdings with the SEC.

B Institutional Holdings

In the US, institutional investment managers holding over \$100M in designated 13F securities must report their quarterly holdings in Form 13F with the Securities and Exchange Commission (SEC). We obtain the data on 13F institutional stock holdings from the Refinitiv Ownership database. Based on Refinitiv's investor labels, we group institutions into six types: banks, pension funds, insurance companies, investment advisors, hedge funds, and others.

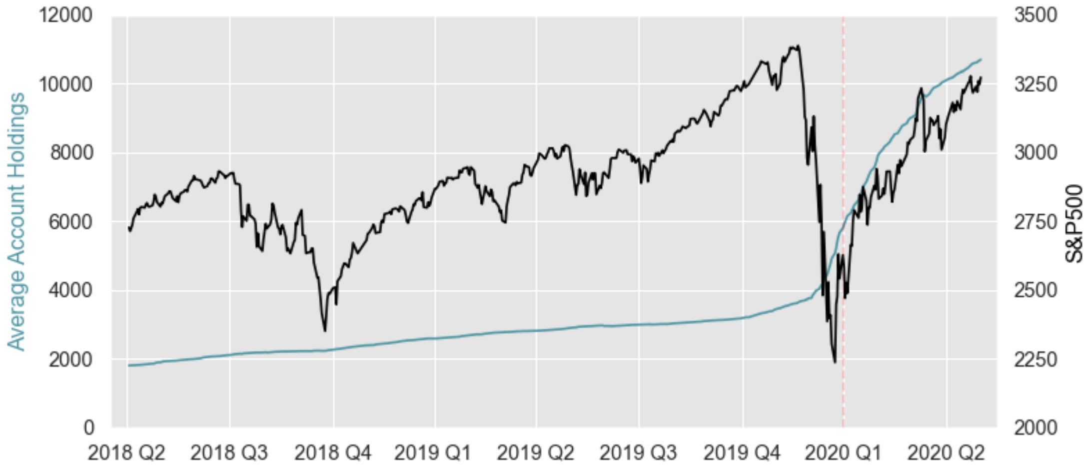
In the construction of institutional portfolios, we closely follow the methodology proposed by KY (2019). In order to keep the exposition as comprehensible and parsimonious as possible, we refer the reader to the detailed description in their paper and to Appendix B of this paper. We compute the percentage of ownership as the number of shares held divided by shares outstanding. Finally, we construct each institution's equity portfolio weights as the dollar holdings in each stock (price times shares held) divided by their AUM. We compute assets under management as the sum of an institution's dollar holdings.

C Robinhood and Household Holdings

Founded in 2013, "Robinhood Markets Inc." has been one of the fastest growing online brokerages. While Robinhood does not disclose the actual number of shares held by its user base, the number of registered accounts invested in a particular stock over time is publicly available and has been compiled on Robintrack from March 2018 to August 2020.[8] We start from the Robintrack dataset and exclude holdings outside our investment universe. Let $H_t(n)$ denote the number of Robinhood accounts holding stock n at time t. Figure 3.1 plots the average number of Robinhood accounts holding a stock, i.e. $\frac{1}{N}\sum_{n=1}^{N} H_t(n)$, from 2018 until 2020. The figure emphasizes the surge in retail trading activity during the onset of the pandemic, a time at which the S&P500 plunged.

[8] The website https://robintrack.net/ relied on Robinhood's own API to obtain the counts of open stock-account positions on the platform. Since Robinhood does not allow its users to short sell, this number solely reflects long positions. Options are also excluded from the count. The API was decommissioned in August 2020.

Figure 3.1: **Evolution of Robinhood Account Holdings**
The left axis of the figure plots the daily average number of Robinhood accounts holding a stock within our investment universe. The number of accounts holding a stock $H_t(n)$ is directly reported on Robinhood and has been compiled until August 2020 by Robintrack. The right axis plots the level of the S&P500 over time.



Appendix C reports further summary statistics on $H_t(n)$ for Q1 and Q2. Between December 2019 and June 2020 the number of open stock-account positions on Robinhood, i.e. $\sum_{n=1}^{N} H_t(n)$, grew from 6 million to 19 million. On May 4$^{\text{th}}$, 2020, Robinhood announced that it currently had 13 million users and had signed 3 million new users in the first quarter alone. These numbers suggest that the boom in retail trading manifested itself in both the entry of new retail traders and an increase in their average portfolio size. The cross-sectional distribution of $H_t(n)$ is strongly (positively) skewed. As of June 2020, the majority of stocks are held by less than 1,000 Robinhood accounts, while some exceptionally popular stocks such as Ford or General Electric are held by more than 500,000 users.

While the construction of institutional portfolios via 13F filings requires little additional assumptions, the construction of the Robinhood portfolio is less straight-forward, as we do not observe the actual amount of shares held by Robinhood traders. However, we can construct a proxy for their portfolio weights using $H_t(n)$. Let $\mathbb{1}_{a,t}(n)$ be an indicator function equal to 1, if Robinhood account a holds stock n in his portfolio at time t. The number of accounts holding n at time t is then given by $H_t(n) = \sum_{a=1}^{U_t} \mathbb{1}_{a,t}(n)$ where U_t is the total number of registered Robinhood users. Cross-sectional differences in $H_t(n)$ indicate which stocks are relatively preferred by Robinhood traders. Assuming that every account holding a stock ($\mathbb{1}_{a,t}(n) = 1$) represents an equal amount of dollars, the representative Robinhood portfolio weight in stock n at time t is given by

$$w_t^{\text{RH}}(n) = \frac{H_t(n)}{\sum_{n=1}^{N} H_t(n)}. \tag{3.1}$$

See Appendix C for a proof as well as a set of alternative and less restrictive assumptions under which (3.1) holds.

In anticipation of its initial public offering on July 29th, 2021, Robinhood released its SEC S-1 filing on July 1st. The file contains the exact number for Robinhood's users and assets under management for March 31st and December 31st of 2020. The number of Robinhood's total users grew from 7.2 to 12.5 million. Over the same horizon, the total equities held by all users quadrupled from \$13.5 to \$53 billion. Using the AUM figures along with the portfolio approximation (3.1), we can reverse-engineer Robinhood traders' share holdings (see Appendix C for details). In order to ensure market clearing, we follow KY (2019) and construct a household sector as the residual shares outstanding not held by 13F institutions and Robinhood traders. In addition, we merge institutions with less than \$10 million under management or without any holdings in the inside and/or outside assets to the household sector, which therefore includes households, small asset managers, and other non-13F institutions.

We are aware of the fact that our approximation of Robinhood traders' portfolio weights critically affects our equilibrium price impact estimates. For example, (3.1) may understate the weight in the largest and overstate the weight in the smallest stocks, if wealthier Robinhood accounts tilt towards market weights. We therefore simulate alternative counterfactual equilibria for deviations from (3.1). In particular, we use market weights $w_t^m(n)$ to construct a pseudo value-weighted Robinhood portfolio $w_t^{\mathrm{RH},v}(n)$ using $\tilde{H}_t(n) = H_t(n)w_t^m(n)$ instead of $H_t(n)$ in (3.1). While Robinhood traders' impact on the equity market in 2020 varies for different portfolio approximations, our main results remain qualitatively unchanged: Given their small market share, Robinhood traders' impact on the valuation of small stocks is large.

3.3 Motivating Facts

In this paper, we explicitly quantify the impact of Robinhood demand during the first half of 2020 and show that, despite their negligible wealth, their *liquidity provision* in Q1 and *recovery amplification* in Q2 was sizeable. We argue that the significance of Robinhood demand is driven by the inelastic nature of financial institutions during the Covid-19 pandemic. Before quantifying these effects in a structural model, we study in reduced form the relationship between Robinhood demand and stock returns during Q1 and Q2. Note that we do not draw any causal implications from these regressions. Our goal is merely to show that stocks more heavily purchased by Robinhood traders exhibit higher returns in Q1 and Q2.

We first confirm this relationship using a cross-sectional regression of quarterly log returns on institutional and Robinhood demand,

$$r_t(n) = b_{0,t} + b_{1,t}\Delta_t^{IO}(n) + b_{2,t}\Delta_t^{RH}(n) + b_{3,t}X_{t-1}(n) + \epsilon_t(n) \quad \text{for} \quad t = \text{Q1, Q2}, \tag{3.2}$$

where $\Delta_t^{IO}(n)$ is the change in institutional ownership for firm n in quarter t, and $\Delta_t^{RH}(n) = H_t(n) - H_{t-1}(n)$ is the quarterly change in the number of Robinhood accounts holding stock n.

$X_{t-1}(n)$ is a vector of lagged controls that includes the stock-specific characteristics described in Section A, the one-period lagged return, and the lagged cash-to-asset ratio. Since we expect this effect to be stronger for small cap companies, we split our sample into size quintiles and estimate (3.2) for each quintile in Q1 and Q2. Panel (a) in Figure 3.2 reports the coefficients on Robinhood demand $b_{2,t}$ across size quintiles. Accounting for known return predictors, high demand stocks belonging to the two smallest size quintiles earned higher returns in both Q1 and Q2. On average, each additional 1,000 invested accounts coincide with 1.95 (1.05) percentage points higher returns in Q1 (Q2) for the smallest size quintile.[9]

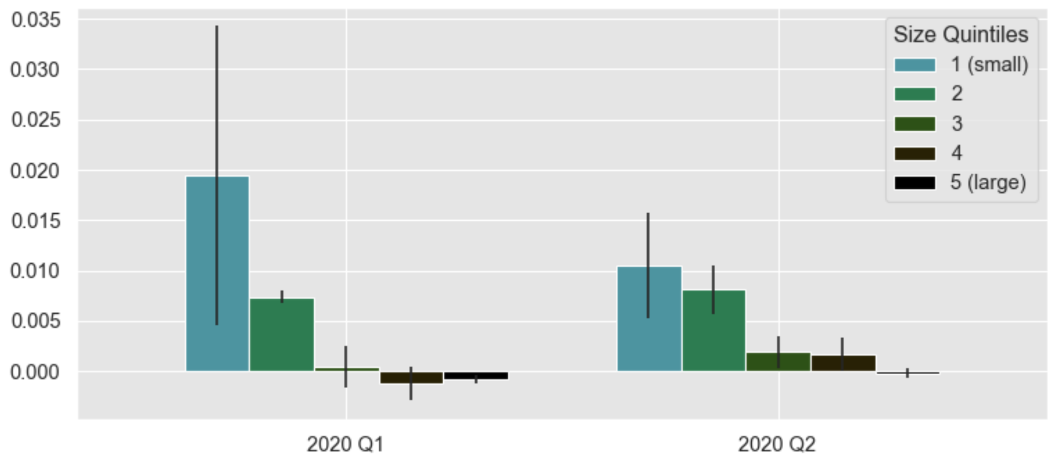
Second, we conduct an event study and compute the abnormal returns of a "Robinhood portfolio" (henceforth RHP). To this end, we sort stocks into deciles based on $\Delta_{Q2}^{RH}(n)$. We then build two long-short portfolios and compute their monthly returns. The first portfolio longs firms in the top decile and shorts firms in the bottom one. The second portfolio also longs the top decile but shorts a set of matched stocks, where matching is based on industry and size. We assess the performance of RHP by analysing the residuals from the Fama-French five-factor model. Factor coefficients are estimated on a 36-months estimation window starting April 2017 and ending March 2021. Panel (b) in Figure 3.2 shows the monthly abnormal returns of RHP, which are significantly greater than zero for the months of April, May, and June.[10] With respect to their past performance, stocks subject to high demand from Robinhood traders in Q2 perform significantly better.

[9] The average change in invested Robinhood accounts $\Delta_t^{RH}(n)$ in the smallest quintile is 988 (1,905) in Q1 (Q2).

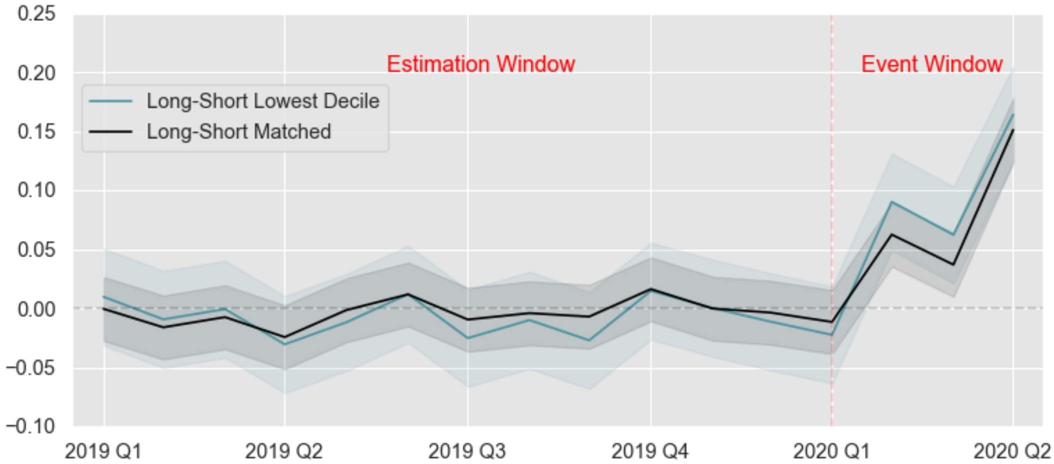
[10] We have conducted the same analysis using deciles based on $\Delta_{Q1}^{RH}(n)$. The abnormal returns display similar patterns, and conclusions remain unchanged. Using quintiles instead of deciles also yields similar results.

Figure 3.2: **Robinhood Demand and Stock Returns in Reduced Form**
Panel (a) reports coefficients on $\Delta_t^{RH}(n)$ from estimating (3.2) for each size quintile in both Q1 and Q2. Coefficients are scaled to reflect changes in thousands. All specifications include industry fixed effects and a set of control variables reported in the main text. Standard errors are clustered at the industry level. Error bars indicate the 95% confidence interval. Panel (b) plots the abnormal returns on the Robinhood long-short portfolios (RHP). Portfolios are equally weighted and rebalanced monthly. The blue line corresponds to the portfolio that longs the decile of stocks with the highest Robinhood demand over Q2 and shorts the bottom decile. The black line corresponds to the long-short portfolio that longs the same top decile and shorts matched stocks from the same industry and of similar size. Abnormal returns are defined in excess of the Fama-French five-factor model estimated on a 36-month window from April 2017 to March 2020. Standard errors are adjusted for autocorrelation using a 12 lag Newey-West correction. The shaded areas indicate 95% confidence bands.

(a) Impact of 1,000 Additional Invested Robinhood Accounts on Returns by Size Quintile



(b) 5-Factor Abnormal Returns



These observations entertain the possibility of Robinhood traders affecting the cross-section of stock returns and the overall recovery process during the Covid-19 pandemic. They do not, however, suffice to conclude that the relationship is causal and remain silent on the role of financial institutions. A *true* (i.e. causal) counterfactual analysis needs to explicitly address the endogeneity of prices and demand. The next section provides a structural model that is able to elicit the causal role of Robinhood traders by estimating empirical demand curves and decomposing prices into investor-specific demand components.

3.4 The Demand System Approach

The "Demand System Approach to Asset Pricing", developed by KY (2019), constitutes a practical modelling framework that - under some modifications - is ideally suited to assess the causal impact of Robinhood traders on the US equity market during Q1 and Q2. In particular, we deviate from KY's methodology by estimating time-invariant demand curves over the panel of observed portfolio holdings including time fixed effects. Thus we can combine Robinhood holdings, which are available at a daily frequency, with quarterly institutional holdings. In particular, this allows us to jointly match daily prices and Robinhood demand and derive daily equilibrium counterfactuals.

A The Model

Following the notation in KY (2019), there are N financial assets and T time periods indexed by $n = 1,...,N$ and $t = 1,...,T$ respectively. The total market equity of asset n is given by $ME_t(n) = P_t(n)S_t(n)$, where $P_t(n)$ and $S_t(n)$ are the price and shares outstanding of asset n. The assets differ in terms of K asset-specific characteristics $X_t(n) \in \mathbb{R}^K$, which are independent of the price $P_t(n)$ and determined by an exogenous endowment process, and log market equity $me_t(n)$, which is endogenously determined in equilibrium as it depends on the equilibrium price $P_t(n)$.

Investor Demand

There are I institutional investors indexed by $i = 1,...,I$, a representative household sector (HH) and a representative Robinhood trader (RH). All $I+2$ investors $j = \{i\}_{i=1}^I, HH, RH$ are short-sale constrained, have log utility and heterogeneous beliefs about expected returns and covariances, and are subject to heterogenous leverage constraints. KY (2019) show that when both expected excess returns and the factor loadings that drive the covariance matrix are linear functions of the $K+1$ asset-specific characteristics $[X_t(n); me_t(n)]$, characteristics-based demand emerges. In other words, the optimal portfolio weights $w_t^j(n)$ are a function of observable stock-specific characteristics:

$$\forall j: \quad \frac{w_t^j(n)}{w_t^j(0)} = \exp\left\{\theta_t^j + \gamma^j me_t(n) + \sum_{k=1}^K \beta_k^j X_{k,t}(n)\right\}\epsilon_t^j(n), \tag{3.3}$$

where $w_t^j(0)$ is the portfolio weight in the outside asset and θ_t^j represents the time-varying demand for all stocks inside the investment universe relative to the outside asset. The investor-specific demand function (3.3) can be estimated empirically using observable portfolio holdings and lies at the heart of the demand system.

Demand Elasticities

Letting $\mathbf{p}_t, \mathbf{q}_t^j \in \mathbb{R}^N$ denote vectors of log prices and log shares held by investor j respectively, it can be shown (see Appendix A) that (3.3) implies the following demand elasticity with respect to the price

$$-\frac{\partial \mathbf{q}_t^j}{\partial \mathbf{p}_t'} = \mathbf{I} - \gamma^j \operatorname{diag}(\mathbf{w}_t^j)^{-1} \mathbf{G}_t^j, \tag{3.4}$$

where $\mathbf{G}_t^j = \operatorname{diag}(\mathbf{w}_t^j) - \mathbf{w}_t^j \mathbf{w}_t^{j\prime}$. Thus γ^j, the demand coefficient associated with market equity, determines the price elasticity of demand. The higher an investor's price elasticity of demand (i.e. the smaller γ^j), the more liquidity this investor provides to the market by accommodating the demand shocks of other investors. In this light, the set of price elasticities $\{\{\gamma^i\}_{i=1}^I, \gamma^{\text{HH}}, \gamma^{\text{RH}}\}$ critically determines the transmission of retail demand shocks to the cross-section of asset prices. If γ^j is 1 for all investors, i.e. the equity market is perfectly inelastic, then retail demand shocks have an infinite impact on the price. Lastly, note that if j holds a diversified portfolio, then $\mathbf{w}_t^j \mathbf{w}_t^{j\prime}$ is small and the diagonal elements of $-\frac{\partial \mathbf{q}_t^j}{\partial \mathbf{p}_t'}$ are approximately $-\frac{\partial q_t^j(n)}{\partial p_t(n)} = 1 - \gamma_j$.

Market Clearing

The system of $I+2$ demand functions given by (3.3) is linked via the market clearing condition for each stock n, which implies that $ME_t(n) = \sum_j A_t^j w_t^j(n)$ for each n, where A_t^j is the AUM of investor j. Plugging in (3.3) and rewriting in vector notation, the equilibrium market clearing prices $\mathbf{p}_t$ solve

$$\mathbf{p}_t = \log\left(\sum_j A_t^j \mathbf{w}_t^j(\mathbf{p}_t)\right) - \mathbf{s}_t, \tag{3.5}$$

where $\mathbf{s}_t \in \mathbb{R}^N$ is the vector of log shares outstanding. The dependence of $\mathbf{w}_t^j(\mathbf{p}_t)$ on $\mathbf{p}_t$ reveals how the explicit estimation of a demand function for each investor j allows to address the interplay of demand and equilibrium prices. When changing e.g. Robinhood traders' demand, (3.5) explicitly accounts for the endogenous demand responses by all other investors. In other words, Robinhood traders' demand affects the portfolio choice of all other investors because equilibrium prices enter their demand functions via γ^j.

B Estimating the Demand System

This section describes how we estimate the demand curves given by (3.3) for institutions, households and Robinhood traders. Note, that while Robinhood holdings are available at a daily frequency, institutional holdings are obtained from SEC form 13F, which is generally filed quarterly. We hence estimate time-invariant demand curves over the panel of observable holdings for each investor. This allows us to feed the estimated demand system with holdings at different frequencies and obtain counterfactuals at a daily frequency.

Instrument Construction

We follow KY (2019) and assume that the characteristics $X_t(n)$ are exogenous to latent demand $\epsilon_t^j(n)$. Prices, however, are inherently endogenous, as they are a function of the latent demand $\epsilon_t^j(n)$ of all investors. Estimating (3.3) without the use of investor-specific instruments for $me_t(n)$ therefore leads to biased demand elasticities. We therefore construct investor-specific instruments for market equity $\widehat{me}_t^j(n)$ using heterogeneity in the investment universes across investors. To this end, let $\mathcal{N}_t^j$ denote the set of stocks in j's investment universe. Stocks that are in the investment universe of more investors have a higher exogenous component of demand and therefore a higher exogenous price. This relies on the assumption that investors' investment universes are (at least partly) exogenously determined. KY estimate investor i's investment universe $\mathcal{N}_t^j$ as all stocks held at some point in the past 11 quarters (including current holdings). While they show that this choice set is quite persistent, it remains unclear why an investor should exert exogenous demand pressure on a stock that is in its universe but not currently held. Furthermore, they implicitly assume that any correlation between latent demand ϵ_t^j and the currently held subset of stocks is sufficiently diluted by including all stocks held over the look-back period of 3 years. To mitigate these concerns, we define $\widehat{\mathcal{N}_t^j} \subset \mathcal{N}_t^j$ as the part of investor j's universe that is exogenously determined. We estimate $\widehat{\mathcal{N}_t^j}$ as the set of stocks that are currently held but were also held at *every* quarter over the past two years. The instrument is equal to the counterfactual market equity obtained, if all investors held equal weighted portfolios across their *exogenous* investment universe. Formally, the instrument is given by

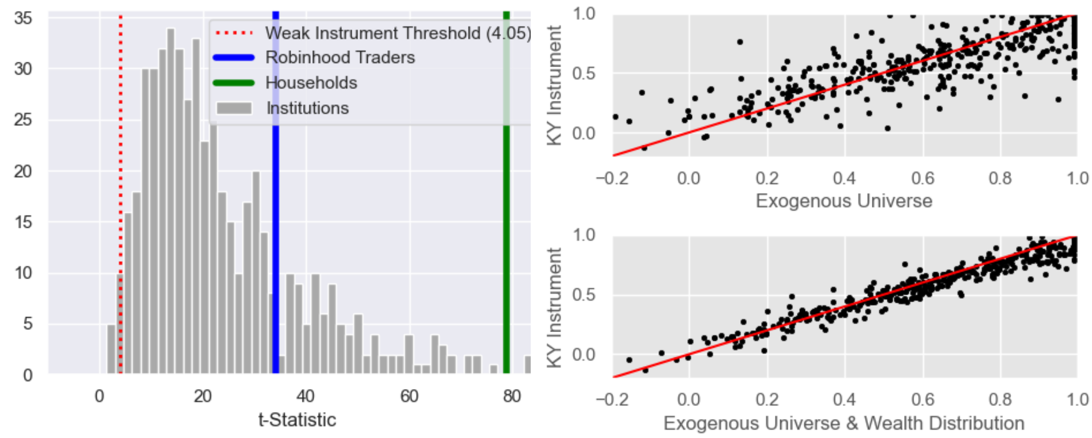
$$\widehat{me}_t^j(n) = \sum_{i \neq j, \text{RH}, \text{HH}} A_t^i \frac{\mathbb{1}_t^i(n)}{1 + |\widehat{\mathcal{N}_{t,i}}|}, \tag{3.6}$$

where $\mathbb{1}_t^i(n)$ is an indicator equal to 1 if stock n is in i's exogenous part of the investment universe $\widehat{\mathcal{N}_{t,i}}$. We assess the instrument's relevance for each investor in a panel regression of $\widehat{me}_t^j(n)$ onto $me_t(n)$ and the exogenous characteristics $X_t(n)$ including time-fixed effects. Panel (a) of Figure 3.3 plots the first stage t-statistic of the instrument for institutions, households, and Robinhood traders.

Figure 3.3: **Weak Instrument Test**

Panel (a) plots the first stage t-statistic on the instrument $\widehat{me}_t^j(n)$ for institutions, households, and Robinhood traders. The red dotted line indicates the weak instrument threshold of 4.05 (see Stock and Yogo (2002)). Panel (b) plots the estimated γ^j for all investors under the different instrument specifications.

(a) 1st Stage t-Statistic (b) Estimates of γ^j for Different Instruments



We can confidently reject the null hypothesis of weak instruments for households, Robinhood traders and almost all institutions.[11] Lastly, the exogeneity of the instrument relies on a wealth distribution $\{A_t^i\}_{i=1}^I$ that is orthogonal to the latent demand shocks driving portfolio weights. One could argue that latent demand, such as ESG preference shocks, potentially drive both prices and flows across institutional investors. Therefore we construct an exogenous wealth distribution, that would be obtained if assets under management were proportional to the number of stocks held $|\mathcal{N}_t^j|$.[12] If we construct the instrument under this alternative wealth distribution, the resulting estimates for γ^j remain largely unchanged. Panel (b) plots the γ^j estimates for all investors under the different instrument specifications. The estimates all lie in the same ballpark and are strongly positively correlated. We conclude that our results on the inelasticity of institutional investors are robust to the construction of the instrument.

[11] The few institutions with marginally insignificant first stage coefficients have a combined AUM of \$30 billion and are pooled together with the household sector.

[12] Formally, the instrument then becomes

$$\widehat{me}_t^j(n) = \sum_{i \neq j, \mathrm{RH}, \mathrm{HH}} |\mathcal{N}_t^i| \frac{\mathbb{1}_t^i(n)}{1 + |\widehat{\mathcal{N}}_{t,i}|}$$

Panel Estimation

We estimate demand (3.3) by constrained generalized method of moments (GMM) under the following moment condition

$$\forall j: \quad \mathrm{E}\left[\log(\epsilon_t^j(n)) \,\middle|\, \widehat{me}_t^j(n), X_t(n)\right] = 0 \quad s.t. \quad \gamma^j < 1. \tag{3.7}$$

The constraint $\gamma^j < 1$ ensures a unique equilibrium price by imposing downward sloping demand for all investors (see KY 2019, proposition 2). We follow KY (2019) and choose log book equity, profitability, investment, market beta, and dividends to book ratio as exogenous characteristics $X_t(n)$. For Robinhood traders, we estimate demand over the panel from March 2018 to August 2020. For institutions and households (the residual party), we have a greater time span of available portfolio holdings and estimate demand over the panel from January 2005 to August 2020. Many 13F institutions hold too concentrated portfolios to consistently estimate their demand. We therefore pool investors with insufficient observations by their institutional type, investment style and size (see Appendix B for details). We emphasize the panel nature of the demand function (3.3), which in contrast to KY (2019) features time-independent demand coefficients γ^j and β_k^j. The purely cross-sectional specification proposed by KY (2019) inherently features time-varying demand curves. Taking logs on both sides of (3.3), we reformulate the demand function into a panel estimation problem for each investor j including time fixed effects. Estimating demand over the panel has a distinct advantage over the quarterly re-estimation in KY (2019). The increased number of observations leads to less pooling across investors and more precisely estimated demand curves. In Appendix D, we re-estimate investor's demand every quarter. While the quarterly estimated coefficients are fairly stable over time, their statistical significance strongly declines.

Estimation Results

Panel (a) of Table 3.1 summarizes the estimated demand curves. For illustrative purposes, we group 13F investors - which comprise over 3,500 institutions - by their respective institutional types: Hedge funds, banks, pensions funds, investment advisors (including mutual funds), insurance companies, and other. We report the AUM-weighted average demand coefficients across investors within a group. Note that, from the estimated demand coefficients γ_j, we can infer each investors' demand elasticity with respect to the price via (3.4). Hedge funds are the most elastic institutional group. Their demand elasticity is 0.46, implying that (ceteris paribus) a 10% price increase in stock n causes hedge funds to sell 5% of their shares held in n. However, hedge funds' ability to provide elasticity to the aggregate equity market is strongly limited, as their holdings account for merely 2% of total holdings within our sample. Pension funds, who are primarily passive long-term investors, have the most inelastic demand. 60% of the US equity market is managed by investment advisors, banks and pension funds, all of whom respond extremely inelastically to price changes. The inelastic nature of the institutional investment management industry is the main driver of micro-inelastic markets and allows the small retail sector to have significant price effects. Households are on average more elastic than institutions and hence provide more liquidity to the market. As of 2020, households hold roughly one third of the US equity market as compared to 60% in the year 2000. Their steadily declining market share in favor of passive institutions may be one reason for the arguable lack of market elasticity observed in recent years. Lastly, Robinhood traders have the most elastic demand, implying that they react most strongly to price changes.

The coefficient on book equity β^j_{be} captures demand for size. While investment advisors and banks have a preference for larger high-dividend stocks, hedge funds tilt towards smaller low-dividend stocks. Robinhood traders have a strong preference for small stocks, which is reflected in their outsized price impact on the smallest quintile of US stocks (see Section B). Demand for market risk β^j_{beta} is lowest for liability constrained institutions such as insurance companies, banks and pension funds, and highest for Robinhood traders. Unsurprisingly, the demand coefficients of hedge funds and pension funds have opposite signs for all exogenous characteristics. Note, furthermore, that the demand functions of households and Robinhood traders strongly diverge. This difference can be attributed to the design of the trading platform and the display of information on Robinhood (see Barber et al. (2020)).

In order to emphasize the inelasticity of institutional demand, Panel (b) groups institutional investors by their elasticity. The groups are defined as follows: *Elastic investors* with $\gamma^j < 0$, *inelastic investors* with $0 \leq \gamma^j < 1$, and *passive investors* with $\gamma^j = 1$. Elastic investors respond to a 1% (non-fundamental) price increase by selling over 1% of their shares. Inelastic investors sell less than 1%, whereas passive investors hold the market portfolio and do not respond to price changes.[13] Less than 1% of the US equity market is managed by elastic institutions with a demand elasticity above 1. The institutional investment industry is largely comprised of passive investors, who hold roughly 40% of the aggregate US market capitalization. Passive investors value-weight their portfolios within their investment universe and hence have a demand elasticity of 0. Without in- and outflows, their demand (in terms of shares held q^j_t) is constant. The large market share of passive investors has strong implications for the micro-elasticity of equity markets and is a key contributor to the significant price impact of small retail investors.

3.5 The Impact of Robinhood Traders on the Equity Market

Having estimated each investor's demand curve, we can express equilibrium prices via the market clearing condition (3.5) as a function of supply and demand components. The market clearing condition is a fixed point problem and allows computing the counterfactual equilib-

[13] The demand responses are approximations based on $-\frac{\partial q^j_t(n)}{\partial p_t(n)} = 1 - \gamma_j$. See Section AA for details.

Table 3.1: **Estimated Demand Functions**
The table reports the estimated demand coefficients for Robinhood traders, households, and 13F institutions. The disaggregated nature of the 13F data allows estimating a separate demand function for each institution. In Panel (a) we group 13F institutions into 6 groups based on Refinitiv's type labels (hedge funds, banks, insurance companies, investment advisors, pension funds) and report AUM-weighted average for each group. In Panel (b) we group of 13F investors into: elastic institutions ($\gamma^j < 0$), inelastic institutions ($0 \leq \gamma^j < 1$) and passive institutions ($\gamma^j = 1$).

(a) Demand Functions

	Wealth Share	Estimated Coefficients						Demand Elasticity
		γ^j	β^j_{be}	$\beta^j_{\text{inv.}}$	β^j_{profit}	β^j_{beta}	$\beta^j_{\text{div/be}}$	
Institutional Investors								
Hedge Funds	2%	0.55	-4.84	8.83	0.31	7.56	-0.43	**0.46**
Insurance Companies	2%	0.71	0.95	5.24	-2.03	-8.41	-1.24	**0.34**
Other	1%	0.85	26.81	3.49	0.69	1.75	0.22	**0.17**
Investment Advisors	50%	0.85	15.17	3.21	0.11	-0.13	0.12	**0.16**
Banks	7%	0.92	18.36	1.13	-0.26	-5.84	0.19	**0.09**
Pension Funds	3%	0.95	4.63	-1.55	-0.02	-2.29	0.03	**0.07**
Households	35%	0.73	7.68	-1.29	0.66	-12.76	0.21	**0.28**
Robinhood Traders	0.2%	0.31	-27.57	5.61	-1.53	26.40	-0.55	**0.69**

(b) Grouping Institutions by Elasticity

Institutional Investors								
Elastic demand	0.07%	-0.06	2.54	9.15	1.71	3.67	-0.82	**1.06**
Inelastic demand	26%	0.63	0.93	3.87	-0.59	-3.07	-0.07	**0.39**
Passive	39%	1	22.92	2.46	0.44	0.89	0.16	**0.02**

rium prices that would clear equity markets under alternative inputs.[14] For example, we can assess the counterfactual equilibrium obtained without the presence of Robinhood traders by setting $A_t^{\text{RH}} = 0$ and redistributing their wealth across all other investors. The counterfactual equilibrium price vector $\mathbf{p}_t \in \mathbb{R}^N$ is an implicit function $\mathbf{g}(.)$ of exogenous characteristics and shares outstanding, as well as the AUM and demand shocks of all investors,

$$\mathbf{p}_t = \mathbf{g}\Bigg(\underbrace{\mathbf{X}_t, \mathbf{s}_t}_{\text{Supply}}, \underbrace{\{A_t^j, \boldsymbol{\epsilon}_t^j\}_{j=\text{I,HH,RH}}}_{\text{Demand}}\Bigg). \tag{3.8}$$

The following sections use (3.8) in three counterfactual experiments to quantify the effect of Robinhood demand on equity markets in 2020.

[14]We numerically solve for counterfactual equilibria using the algorithm proposed in KY (2019). Uniqueness is guaranteed by the coefficient constraint that $\gamma^j < 1$ for all investors.

A Variance Decomposition

We first use (3.8) to decompose the cross-section of stock returns during the first two quarters of 2020 into supply- and demand-driven components. We label changes in the exogenous characteristics $\mathbf{X}_t$ and log shares outstanding $\mathbf{s}_t$ as *supply* changes. *Demand* changes refer to changes in assets under management $(A^{\mathrm{I}}_t, A^{\mathrm{HH}}_t, A^{\mathrm{RH}}_t)$ and changes in latent demand $(\boldsymbol{\epsilon}^{\mathrm{I}}_t, \boldsymbol{\epsilon}^{\mathrm{HH}}_t, \boldsymbol{\epsilon}^{\mathrm{RH}}_t)$. To reduce the notational clutter, we define the set of supply and demand variables as $\mathbf{S}_t = (\mathbf{X}_t, \mathbf{s}_t)$ and $\mathbf{D}^j_t = (A^j_t, \boldsymbol{\epsilon}^j_t)$ respectively. By iteratively changing supply and investor-specific demand to their next period's values, we can decompose the vector of log returns $\mathbf{r}_{t+1} = \mathbf{p}_{t+1} - \mathbf{p}_t$ into its subcomponents,

$$
\begin{aligned}
\mathbf{r}^{\mathrm{S}}_{t+1} &= \mathbf{g}\big(\mathbf{S}_{t+1}, \mathbf{D}^{\mathrm{I}}_t, \mathbf{D}^{\mathrm{HH}}_t, \mathbf{D}^{\mathrm{RH}}_t\big) - \mathbf{p}_t,\\
\mathbf{r}^{\mathrm{I}}_{t+1} &= \mathbf{g}\big(\mathbf{S}_{t+1}, \mathbf{D}^{\mathrm{I}}_{t+1}, \mathbf{D}^{\mathrm{HH}}_t, \mathbf{D}^{\mathrm{RH}}_t\big) - \mathbf{g}\big(\mathbf{S}_{t+1}, \mathbf{D}^{\mathrm{I}}_t, \mathbf{D}^{\mathrm{HH}}_t, \mathbf{D}^{\mathrm{RH}}_t\big),\\
\mathbf{r}^{\mathrm{HH}}_{t+1} &= \mathbf{g}\big(\mathbf{S}_{t+1}, \mathbf{D}^{\mathrm{I}}_{t+1}, \mathbf{D}^{\mathrm{HH}}_{t+1}, \mathbf{D}^{\mathrm{RH}}_t\big) - \mathbf{g}\big(\mathbf{S}_{t+1}, \mathbf{D}^{\mathrm{I}}_{t+1}, \mathbf{D}^{\mathrm{HH}}_t, \mathbf{D}^{\mathrm{RH}}_t\big),\\
\mathbf{r}^{\mathrm{RH}}_{t+1} &= \mathbf{p}_{t+1} - \mathbf{g}\big(\mathbf{S}_{t+1}, \mathbf{D}^{\mathrm{I}}_{t+1}, \mathbf{D}^{\mathrm{HH}}_{t+1}, \mathbf{D}^{\mathrm{RH}}_t\big).
\end{aligned}
$$

Note that $\mathbf{g}(.)$ is concave in investors' dollar demand. Thus we purposely place the change in Robinhood traders' demand *after* having accounted for institutions' and households' demand change in order to leave the least scope for Robinhood traders to affect returns. Using the fact that $\mathbf{r}_{t+1} = \mathbf{r}^{\mathrm{S}}_{t+1} + \mathbf{r}^{\mathrm{I}}_{t+1} + \mathbf{r}^{\mathrm{HH}}_{t+1} + \mathbf{r}^{\mathrm{RH}}_{t+1}$, we decompose the cross-section of returns during the pandemic into its underlying supply and demand components[15],

$$
\mathrm{Var}(\mathbf{r}_{t+1}) = \mathrm{Cov}(\mathbf{r}_{t+1}, \mathbf{r}^{\mathrm{S}}_{t+1}) + \mathrm{Cov}(\mathbf{r}_{t+1}, \mathbf{r}^{\mathrm{I}}_{t+1}) + \mathrm{Cov}(\mathbf{r}_{t+1}, \mathbf{r}^{\mathrm{HH}}_{t+1}) + \mathrm{Cov}(\mathbf{r}_{t+1}, \mathbf{r}^{\mathrm{RH}}_{t+1}). \tag{3.9}
$$

The structural nature of the decomposition explicitly takes into account the endogeneity of prices and demand, and therefore allows for a quantification of Robinhood traders' contribution to the cross-section of stock returns in Q1 and Q2 that goes far beyond simple regressions of returns onto holdings as in e.g. Ozik et al. (2020) or Glossner et al. (2020).

[15]The variance and covariance terms are purely *cross-sectional*, i.e. they are computed as $\mathrm{Var}(\mathbf{r}_{t+1}) = \frac{1}{N}\sum_{n=1}^{N}\big(r_{t+1}(n) - \bar{r}_{t+1}\big)^2$.

Figure 3.4: **Variance Decomposition in Q1 and Q2**
The figure reports the decomposition of return variation during the first and second quarter of 2020. Cross-sectional variance is decomposed into investor-specific subcomponents, which add up to 1 as given by equation (3.9). The *Supply* bar indicates the variation in returns explained by changes in fundamentals and shares outstanding. The *Demand* bars (split into institutions, households, and Robinhood traders) indicate the variation explained by changes in latent demand and assets under management.

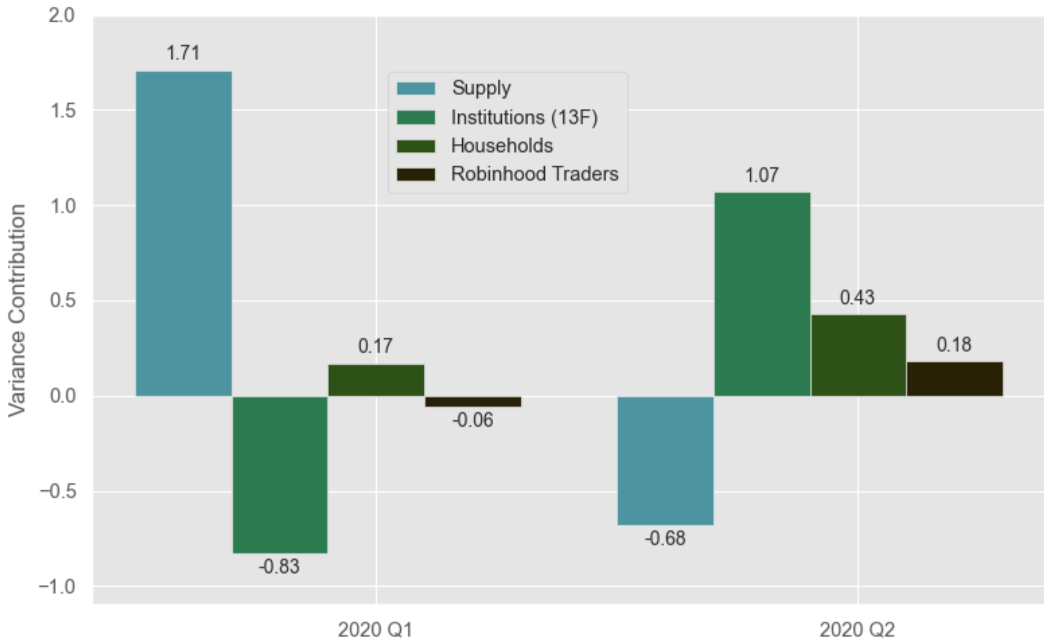


Figure 3.4 reports the decomposition of the cross-sectional variation into changes in supply (i.e. changes in exogenous characteristics and shares outstanding) and demand (i.e. changes in the AUM and latent demand of institutions, households, and Robinhood traders).[16] In Q1, the majority of the cross-sectional variation in returns can be explained by changes in stock-specific characteristics unrelated to price. This is reasonable as the pandemic had real effects on the companies' balance sheets, which enter investor demand via the fundamentals *profitability*, *investment*, and *book equity*. The deterioration of these exogenous characteristics led to lower investor demand across the board and hence lower stock prices. The negative contribution of institutions' changes in AUM and latent demand shows that they provided liquidity by accommodating some of the supply-side induced price pressure. However, relative to their small market share, the liquidity provision by Robinhood traders is much stronger. Despite holding less than 0.2% of the aggregate equity market, their contribution to the cross-sectional variation in stock returns is -11%. This finding is in line with Ozik et al. (2020) and Glossner et al. (2020), who show that Robinhood traders attenuated the rise in illiquidity during the first quarter of 2020.

By contrast, the variations in returns observed during Q2 are purely demand driven. In fact, the

[16]In the Online Appendix, we decompose the variance contribution of institutions into different institutional types (i.e. hedge funds, pension funds, etc.).

counterfactual returns observed without demand effects (i.e. only considering the real effects from changes in fundamentals) are negatively related to observed returns. Most importantly, the demand of Robinhood traders accounts for 10% of the cross-sectional variation in returns observed during the recovery in Q2. Considering their small market share, Robinhood traders' contribution is enormous. Note that this impact is directly related to the inelastic nature of institutional demand. If institutional investors would react more elastically to price changes (e.g. via a greater active asset management industry), the impact of Robinhood traders' demand shocks would be considerably smaller. In the Online Appendix, we report a variance decomposition under the assumption that all investors have a demand elasticity of $-\frac{\partial q_t^j(n)}{\partial p_t(n)} = 1$. Within this more elastic equity market Robinhood traders' impact still remains sizable.

B Counterfactual Equity Market Without Robinhood Traders

The equilibrium pricing function (3.8) allows us to infer the impact of retail trading activity on both the crash in Q1 as well as the recovery in Q2. We construct a counterfactual recovery process for 2020, which would have been observed without the surge in retail trading activity. To this end, note that the market clearing condition allows us to extract the equilibrium price vector $\mathbf{p}_t^{-\text{RH}}$, that clears equity markets in 2020 without Robinhood traders. We take a conservative approach by assuming that without the existence of Robinhood traders, their assets under management were instead invested in the market portfolio. This is consistent with the market clearing condition that investors' wealth shares must sum to 1. Without the existence of the brokerage platform in 2020, the shares traded on Robinhood would be held by all other investors. If we could observe the origins of Robinhood's market share as flows from e.g. investment advisors to retail trading platforms, then we could take a more explicit stance on the redistribution of wealth shares. In an earlier version of the paper we conducted a counterfactual experiment that withdraws Robinhood traders' assets under management from the equity market without redistribution. While the cross-sectional repricing effects remain quantitatively unchanged, the repricing effects for the stock market as a whole are naturally smaller under redistribution. In fact, withdrawing Robinhood traders' AUM leads to a mechanical decline in aggregate valuations, while under redistribution the value of the aggregate equity market market remains roughly unchanged.

Formally, the counterfactual equilibrium price vector without Robinhood traders is given by

$$\mathbf{p}_t^{-\text{RH}} = \mathbf{g}\Big(\mathbf{X}_t, \mathbf{s}_t, \{\gamma^j, \beta^j, \tilde{A}_t^j, \boldsymbol{\epsilon}_t^j\}_{j\neq\text{RH}}\Big), \tag{3.10}$$

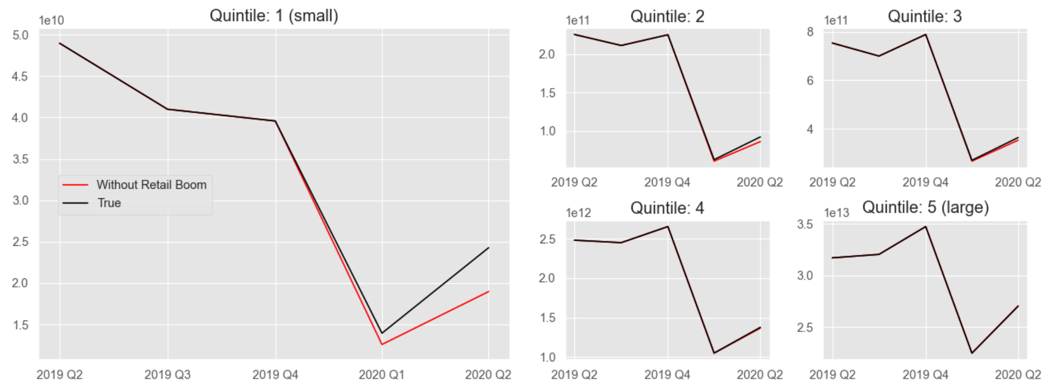
where $\tilde{A}_t^j = A_t^j + A_t^{\text{RH}} \frac{A_t^j}{\sum_{k\neq\text{RH}} A_t^k}$. This is equivalent to assigning to Robinhood traders the wealth-weighted demand curve of all other investors, as in Koijen et al. (2022). Using the counterfactual price vector $\mathbf{p}_t^{-\text{RH}}$, we can gauge the contribution of Robinhood traders to the crash and recovery of a subset of stocks $\mathcal{N}_s \subseteq \mathcal{N}$ by computing the counterfactual aggregate market capitalization

$$M_{t,\mathcal{N}_s}^{-\mathrm{RH}} = \sum_{n \in \mathcal{N}_s} \exp\{p_t^{-\mathrm{RH}}(n) + s_t(n)\}. \tag{3.11}$$

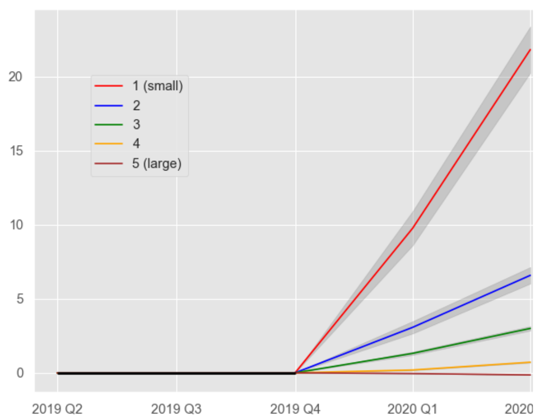
Figure 3.5 reports the observed total market capitalization $M_{t,\mathcal{N}_s}$ and counterfactual market capitalization $M_{t,\mathcal{N}_s}^{-\mathrm{RH}}$ across size quintiles during the Covid-19 pandemic. Panel (a) plots the absolute valuation effects in dollar terms. The black line plots the truly observed market capitalization on a quarterly basis, whereas the red line plots the counterfactual market capitalization in the absence of Robinhood traders. Panel (b) shows the repricing effects as the valuation delta relative to the true market capitalization $\frac{M_{t,\mathcal{N}_S} - M_{t,\mathcal{N}_S}^{-\mathrm{RH}}}{M_{t,\mathcal{N}_S}}$.

Figure 3.5: **Counterfactual US Market Capitalization**
Panel (a) reports the true and counterfactual market capitalization across size quintiles with and without Robinhood traders during the Covid-19 pandemic. The black line illustrates the observed aggregate market capitalization for each size quintile, whereas the red line plots the counterfactual market capitalization in the absence of Robinhood traders. Panel (b) reports the relative repricing of the size quintiles due to Robinhood traders. The shaded areas indicate 99% confidence bands, which we compute via simulating alternative demand curves (see Section D). Panel (c) plots the counterfactual market capitalization of the smallest quintile in the absence of Robinhood traders for different levels of A_{Q2}^{RH} ranging from \$35 to \$95 billion. Size quintiles are computed over all stocks within our investment universe using their market equity as of January 2020.

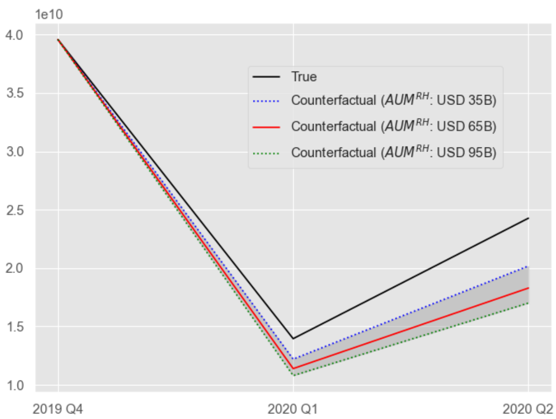
(a) \$ Valuation Effects: Size Quintiles



(b) % Repricing: Size Quintiles



(c) A_t^{RH} Simulation: Smallest Quintile



The impact of Robinhood traders on the valuation of small stocks is substantial. Robinhood traders drove 25% of the valuations of the smallest quintile during Q2. In other words, the aggregate market capitalization of small cap stocks would have been 25% lower without the additional demand from Robinhood traders. Moving to larger size quintiles, their incremental contribution decreases monotonically. In fact, the presence of Robinhood traders has a slightly negative effect on the valuation of the largest quintile of US stocks. Under the redistribution of Robinhood's assets, the counterfactual effects on the aggregate equity market are negligible. This is owed to Robinhood traders' inability to affect the valuation of large cap stocks, which account for the majority of the aggregate market capitalization. Even with an extremely inelastic institutional sector, the retail sector's dollar demand shocks are too small relative to the market capitalization of e.g. the big tech stocks (i.e. Amazon, Apple, Google, Facebook and Microsoft). When moving to deciles instead of quintiles the effects are even more pronounced. Without Robinhood traders, the aggregate market capitalization of the smallest decile of US stocks would have been roughly 40% lower in the first two quarters of 2020.[17] We note that retail traders' impact on the equity market going forward critically depends on their growth in assets under management. In Panel (c) of Figure 3.5, we therefore simulate counterfactual valuations of the smallest quintile under varying levels of Robinhood traders' AUM in Q2. Naturally, Robinhood traders' impact declines with lower AUM estimates. However, even under an AUM estimate as low as $35 billion, their absence would nevertheless lead to a 17% lower aggregate market capitalization of the smallest quintile of US stocks. We therefore confirm that our main finding, the strong repricing of smaller stocks due to retail trading, is likely to continue. Lastly, we assess Robinhood traders' impact at the industry level. The relative valuation effects are strongest for energy and industrial stocks with Robinhood traders increasing aggregate valuations by almost 1%. The effects are substantial given the fact that the aggregate market capitalization of these industries greatly exceed Robinhood's total assets under management.

Overall, the retail investment boom led to a considerable dampening of the price deterioration of small stocks during the first quarter of 2020. Thus we quantitatively confirm the hypothesis made by Glossner et al. (2020) and Ozik et al. (2020) that Robinhood traders provided liquidity in Q1. The contribution of the retail crowd during the recovery in Q2 is arguably even more interesting, as it encapsulates the key point that this paper is trying to highlight: The micro-inelastic nature of financial markets - driven by the large market share of passive institutions - allows the relatively small retail sector to have substantial aggregate pricing effects. The surprisingly large aggregate impact of the small retail sector is owed to the fact that a large fraction of the equity market is managed by institutional investors with price-inelastic demand curves.

C Do Robinhood Traders Increase Volatility?

A large empirical literature on the interaction of stock returns and retail trading activity documents short-term contrarian behaviour by individual investors (see e.g. Kaniel et al. (2008) and Barber et al. (2008) and the references therein). More recently, Peress and Schmidt (2020) find that transitory declines in the intensity of noise trading lead to deteriorating liquidity and higher volatility.[18] With a steadily growing retail investment crowd, quantifying

[17]Blume et al. (2014) find that, in recent times, institutional investors have overweighted small stocks. This disproportionate presence of inelastic institutions can further explain how Robinhood traders' demand shocks are amplified beyond their small dollar value.

[18]Using a french reform that permanently discouraged retail trading for a subset of stocks, Foucault et al. (2011) find that reduced retail trading activity improves liquidity and decreases volatility. Similary, Eaton et al. (2021) find

the impact of retail trading at a higher frequency is potentially important. Our structural model allows for a quantification of the impact of Robinhood demand on stock-specific volatility, which would be difficult to do in a reduced-form framework.

Note, that the estimated demand system based on quarterly data is not ideally suited to assess the time-series volatility of individual stocks (as opposed to the cross-sectional variation in Section A). We therefore extend the demand system approach to jointly match holdings and prices at a daily frequency. Unfortunately, Robinhood traders are the only market participants for which we obtain stock holdings at a daily frequency. The availability of institutional holdings at a quarterly frequency makes it impossible to account for the micro-structural interplay between institutional and retail trading).[19] We can nevertheless compute daily counterfactual prices obtained if Robinhood traders were the only active entity, i.e. under the absence of fundamental changes and institutional trading.

To this end, let t_d denote the d^{th} day in quarter t and t_0 the beginning-of-quarter date. During a quarter, we set stock-specific fundamentals and institutional demand to their t_0 values. We first infer daily Robinhood demand shocks as

$$\epsilon^{\text{RH}}_{t_d}(n) = \log\left(\frac{w^{\text{RH}}_{t_d}(n)}{w^{\text{RH}}_{t_d}(0)}\right) - \gamma^{RH} me_{t_d}(n) - \sum_{k=1}^{K} \beta^{\text{RH}}_k X_{k,t_0}. \tag{3.12}$$

We also estimate the daily evolution of Robinhood's AUM $A^{\text{RH}}_{t_d}$ by fitting a polynomial to the evolution of account holdings (see 3.1) that matches the quarterly AUM figures reported in Robinhood's S1 filing. For each trading day within a quarter, we then set Robinhood traders' latent demand and their assets under management to their daily values and compute the counterfactual daily return obtained due to Robinhood demand only. Formally, the counterfactual daily return due to Robinhood trading only is given by

$$\mathbf{r}^{\text{RH}}_{t_{d+1}} = \mathbf{p}^{\text{RH}}_{t_{d+1}} - \mathbf{p}^{\text{RH}}_{t_d} \tag{3.13}$$

where

$$\mathbf{p}^{\text{RH}}_{t_d} = \mathbf{g}\Big(\mathbf{X}_{t_0}, \mathbf{s}_{t_0}, \{A^j_{t_0}, \boldsymbol{\epsilon}^j_{t_0}\}_{j\neq\text{RH}}, A^{\text{RH}}_{t_d}, \boldsymbol{\epsilon}^{\text{RH}}_{t_d}\Big).$$

Daily changes in fundamentals and daily institutional trading (which are both unobservable) account for the difference between daily observed returns $\mathbf{r}_{t_{d+1}}$ and the counterfactual returns due to Robinhood trading only. Note that the equilibrium return observed without Robinhood trading can hence be written as $\mathbf{r}^{\text{-RH}}_{t_{d+1}} = \mathbf{r}_{t_{d+1}} - \mathbf{r}^{\text{RH}}_{t_{d+1}}$. Based on this decomposition we define the

that intense herding by retail traders may lower market quality.

[19] Our structural model therefore cannot quantify and distinguish between e.g. adverse selection risk (as in Kyle (1985)) and inventory risk (as in Grossman and Miller (1988)).

relative increase in stock-specific volatility due to Robinhood trading as

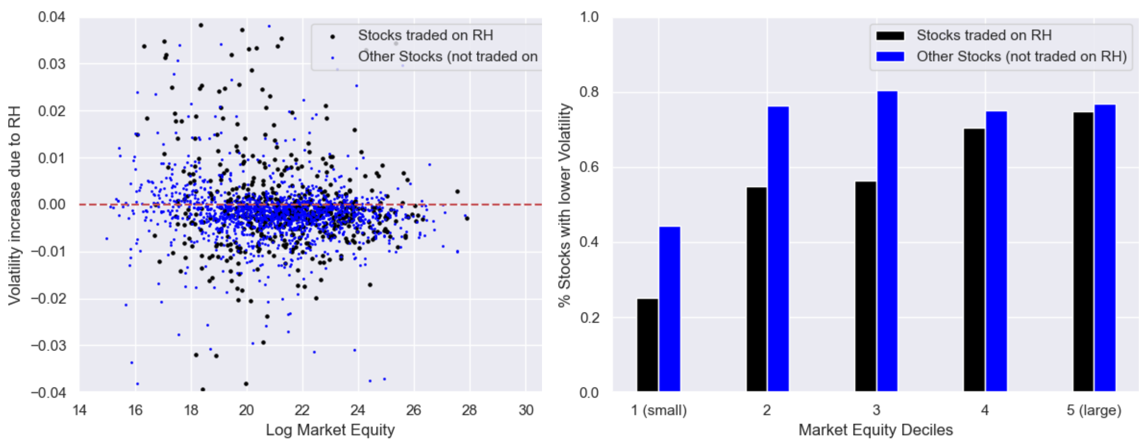
$$\Delta\sigma^{\mathrm{RH}}(n) = \frac{\mathrm{Cov}(\mathbf{r}^{\mathrm{RH}}_{t_{d+1}}, \mathbf{r}_{t_{d+1}})}{\mathrm{Var}(\mathbf{r}_{t_{d+1}})}.^{20} \tag{3.14}$$

We compute $\Delta\sigma^{\mathrm{RH}}(n)$ for each stock n in our sample using true and counterfactual daily returns. If Robinhood traders are contrarian with respect to n, they provide liquidity to the market and hence lower stock-specific volatility as $\mathrm{Cov}(\mathbf{r}^{\mathrm{RH}}_{t_{d+1}}, \mathbf{r}_{t_{d+1}}) < 0$. Note, that a price increase of an individual stock has spill-over effects to all other stocks as market participants rebalance their portfolios by substituting towards other stocks in their investment universe. Thus Robinhood trading affects the volatility of all stocks, even the ones not directly traded on Robinhood.

Figure 3.6: **Impact on Daily Return Volatility**

Panel (a) plots the relative increase in volatility due to Robinhood traders $\Delta\sigma^{\mathrm{RH}}(n) = \frac{\mathrm{Cov}(\mathbf{r}^{\mathrm{RH}}_{t_{d+1}}, \mathbf{r}_{t_{d+1}})}{\mathrm{Var}(\mathbf{r}_{t_{d+1}})}$ for all stocks in the sample relative to their size. There are a few herding stocks (less than 5%), for which $\Delta\sigma^{\mathrm{RH}}(n)$ takes very large values. For illustrational purposes, we limit the y-axis at 4%. Panel (b) plots the fraction of stocks for which Robinhood traders lower volatility (i.e. $\Delta\sigma^{\mathrm{RH}}(n) < 0$) by size quintiles based on market equity.

(a) $\Delta\sigma^{\mathrm{RH}}(n)$ for all stocks (b) Median $\Delta\sigma^{\mathrm{RH}}(n)$ by size quintile



The left panel of Figure 3.6 plots $\Delta\sigma^{\mathrm{RH}}(n)$ for all stocks in the sample, split by whether they are actively traded by Robinhood traders. The presence of Robinhood traders lowers the volatility for over two-thirds of all stocks. This suggests, that the contrarian behaviour of Robinhood traders may improve overall liquidity by providing immediacy to institutional trades. Robinhood traders are less constrained than institutions, whose ability to provide liquidity is limited by e.g. investment mandates and agency conflicts. Notably, 71% of all stocks not directly traded on Robinhood have a lower volatility due to the presence of Robinhood

[20] The absolute increase in volatility due to Robinhood traders is given by $\mathrm{Var}(\mathbf{r}_{t_{d+1}}) - \mathrm{Var}(\mathbf{r}^{\text{-RH}}_{t_{d+1}}) = 2\mathrm{Cov}(\mathbf{r}^{\mathrm{RH}}_{t_{d+1}}, \mathbf{r}_{t_{d+1}}) - \mathrm{Var}(\mathbf{r}^{\mathrm{RH}}_{t_{d+1}})$. Omitting the latter term, which artificially drives up the liquidity provision by Robinhood traders, and scaling by the observed volatility yields (3.14).

investors. This implies that the liquidity provision by Robinhood traders has positive spillover effects to other stocks. The right panel of Figure 3.6 reports the fraction of stocks for which Robinhood traders lower volatility (i.e. $\Delta\sigma^{\text{RH}}(n) < 0$) by size quintile. Only for the smallest quintile, Robinhood trading activity increases volatility for the majority of stocks. These stocks are sufficiently small such that Robinhood trading drives market prices as opposed to providing liquidity to institutional trades. For all other quintiles the contrarian behaviour of Robinhood traders tends to accommodate institutional trades and leads to lower volatility. Note, that demand shocks for large stocks have high cross-sectional spillover effects because they constitute a large portion of institutional portfolios and therefore cause considerable portfolio rebalancing. In contrast, the spillover effects from smaller stocks are negligible. Thus despite Robinhood traders increasing the volatility of small stocks, the overall liquidity spillovers are positive.

D Simulating Alternative Demand Curves

In order to underline the robustness of our results, we simulate the counterfactual equilibria without Robinhood traders under alternative demand curve estimates. This allows us to compute standard errors around our repricing measures. To this end, we estimate the covariance matrix of the coefficients $\Omega^j = \text{Var}(\hat{\beta}^j, \hat{\gamma}^j)$ for each investor, taking into account both heteroskedasticity as well as autocorrelation in the latent demand component (the residuals from the demand estimation). As we have over 1,400 observations for each investor, the joint distribution of the estimated demand coefficients is approximately Gaussian. Using Ω^j and the estimated coefficients from Table 3.1, we simulate 100 alternative demand curves for each investor, i.e.$\{\gamma^j_{(s)}, \beta^j_{(s)}, \boldsymbol{\epsilon}^j_{t,(s)}\}_{s=1}^{100}$. For each set of simulated demand curves, we re-estimate the counterfactual equilibrium without Robinhood traders, which results in 100 simulated counterfactual price vectors $\{\mathbf{p}^{-\text{RH}}_{t,(s)}\}_{s=1}^{100}$. Standard errors around the counterfactual aggregate market equity are computed over the simulated equilibrium samples. Table 3.2 reports the repricing due to Robinhood traders for the aggregate US stock market, size percentiles, and the energy and industrial sector. Note that the Robinhood-induced repricing is highly significant, which is a direct consequence of the low standard errors around the estimated demand coefficients. Low standard errors imply that changes in the demand coefficients over the simulated samples are small, causing minimal changes in the counterfactual price vector and consequently low standard errors around the equilibrium. This gives us confidence that the strong impact of Robinhood traders is not driven by misestimated demand curves.

E Price Impact at the Stock Level

A widespread narrative during Q2 was that some of the large jumps in stock prices are owed to retail investors' demand shocks. The demand system offers an ideal setting to put this narrative through an analytical test. To this end, we use (3.10) to assess each stock's counterfactual return

Table 3.2: **Simulated Counterfactual Repricing**
The table reports the repricing of the market, the size percentiles, and of the energy and industrial sector without the presence of Robinhood traders in Q1 and Q2 respectively. Repricing is defined as $(M_{t,\mathcal{N}_s} - M^{-\text{RH}}_{t,\mathcal{N}_s})/M_{t,\mathcal{N}_s}$. We simulate s=1,...,100 demand curves for each investor from the multivariate normal distribution given by $\mathcal{N}\big((\hat{\beta}^j, \hat{\gamma}^j), \Omega^j\big)$. Given 100 different demand curves for each investor $\{\gamma^j_{(s)}, \beta^j_{(s)}, \boldsymbol{\epsilon}^j_{t,(s)}\}_{s=1}^{100}$, we compute 100 counterfactual equilibrium price vectors $\{\mathbf{p}^{-\text{RH}}_{t,(s)}\}_{s=1}^{100}$. Standard errors are then computed over the sample of 100 simulated counterfactuals.

	Repricing (%)					
	All Stocks	Largest Quintile	Smallest Quintile	Smallest Decile	Energy Stocks	Industrial Stocks
2020 Q1						
Repricing (%)	-0.03	-0.11	18.36	38.93	0.51	0.36
Standard Errors (%)	0.00	0.00	0.44	0.80	0.01	0.01
2020 Q2						
Repricing (%)	-0.03	-0.16	24.94	41.32	0.89	0.83
Standard Error (%)	0.00	0.00	0.60	0.96	0.02	0.02

without the demand shocks from Robinhood traders as $R^{-\text{RH}}_{t+1}(n) = \exp\{p^{-\text{RH}}_{t+1}(n) - p_t(n)\} - 1$. Table 3.3 reports the 10 stocks within the largest decile of US companies in our universe for which Robinhood traders had the greatest impact on returns in Q1 and Q2. Panel (a) and (b) respectively confirm the proposed liquidity and amplification channel of Robinhood demand. Robinhood demand substantially alleviated the negative returns observed in Q1. Ford's return during Q1 would have been over 14 percentage points (pp) lower without the liquidity provision from Robinhood traders. Similarly, General Electric's share price would have dropped by an additional 12 pp without the demand from Robinhood. The return implications are even more pronounced during the recovery in Q2. Ford's return would have been over 40 pp lower, and American Airlines' return would have been over 60 pp lower without the additional demand coming from Robinhood traders. The magnitude of Robinhood demand effects is particularly remarkable given the size of these companies.

In order to better understand the source of the demand effects, we decompose Robinhood traders' price impact into an elasticity and ownership channel. Differentiating (3.5) with respect to the latent demand of Robinhood traders reveals the coliquidity matrix $\frac{\partial \mathbf{p}_t}{\partial \log(\boldsymbol{\epsilon}^{\text{RH}}_t)'} \in \mathbb{R}^{N\times N}$ as in KY (2019, equation 23). Assuming that $\text{diag}(\mathbf{w}_j) - \mathbf{w}_j\mathbf{w}_j' \approx \text{diag}(\mathbf{w}_j)$, the diagonal elements of the coliquidity matrix simplify to

$$\frac{\partial p_t(n)}{\partial \log(\epsilon^{\text{RH}}_t(n))} = \mathcal{M}_t(n) z^{\text{RH}}_t(n), \tag{3.15}$$

where $\mathcal{M}_t(n) = \frac{1}{1-\sum_j z^j_t(n)\gamma^j}$ and $z^j_t(n)$ is the fraction of shares outstanding of n held by j (see Appendix A for a proof). $\mathcal{M}_t(n)$ can be interpreted as a cross-sectional multiplier, which

Table 3.3: **Robinhood traders' price impact on large stocks**

Within the largest decile of US stocks, we choose the 10 stocks with the highest $p_{t+1}^{-\text{RH}}(n) - p_t(n)$ for Q1 and Q2. We report the actually observed return R_{t+1} and the counterfactually observed return without Robinhood demand $R_{t+1}^{-\text{RH}}(n) = \exp\{p_{t+1}^{-\text{RH}}(n) - p_t(n)\} - 1$. The last two columns in each panel decompose the price impact into the multiplier $\mathcal{M}_t(n) = 1/(1 - \sum_j z_t^j(n)\gamma^j)$ and the fraction of shares held by Robinhood traders $z_t^{\text{RH}}(n)$. $R_t, R_t^{-\text{RH}}(n)$, and $z_t^{\text{RH}}(n)$ are reported in %.

(a) 2020 Q1

Company Name	R_{Q1}	$R_{Q1}^{\text{-RH}}$	$\mathcal{M}_{Q1}(n)$	$z_{Q1}^{\text{RH}}(n)$
Ford Motor Co	-48.06	-62.69	5.37	9.24
Carnival Corp	-74.09	-78.65	5.17	5.75
Royal Caribbean Cruises Ltd	-75.90	-79.86	4.88	5.58
Twitter Inc	-23.37	-30.36	6.09	2.48
MGM Resorts International	-64.53	-67.74	5.46	2.85
United Airlines Holdings Inc	-64.18	-67.15	4.00	3.61
Delta Air Lines Inc	-51.21	-54.99	4.44	2.92
General Electric Co	-28.85	-34.17	6.14	2.01
Halliburton Co	-72.01	-74.06	4.45	2.69
Advanced Micro Devices Inc	-0.83	-4.68	6.51	1.02

(b) 2020 Q2

Company Name	R_{Q2}	$R_{Q2}^{\text{-RH}}$	$\mathcal{M}_{Q2}(n)$	$z_{Q2}^{\text{RH}}(n)$
Carnival Corp	24.68	-22.67	5.96	12.54
Ford Motor Co	25.88	-16.41	5.40	11.47
Delta Air Lines Inc	-1.68	-28.12	4.86	9.92
United Airlines Holdings Inc	9.70	-17.02	4.39	10.02
MGM Resorts International	42.37	11.02	5.50	7.32
Royal Caribbean Cruises Ltd	56.36	25.14	4.99	6.79
General Electric Co	-13.98	-26.83	6.04	4.18
Twitter Inc	21.29	8.31	6.14	2.97
Southwest Airlines Co	-4.02	-11.70	4.72	3.06
Halliburton Co	89.49	76.09	4.75	2.53

is inversely related to the market's price elasticity of demand for n. The more elastic the investors in stock n are, the better they accommodate Robinhood traders' demand shocks and the lower $\mathcal{M}_t(n)$ is.[21] In this light, $\mathcal{M}_t$ is the micro-pendant to the macro-multiplier in Gabaix and Koijen (2021). Equation 3.15 highlights the sources of Robinhood traders' price impact on stock n: Either the investors in n are inelastic ($\mathcal{M}_t$ is large), or Robinhood traders hold a substantial fraction of n's shares outstanding (z_t^{RH} is large), or both. The size of the cross-sectional multiplier $\mathcal{M}_t$ helps explain the large price impact of Robinhood demand despite their small ownership share z_t^{RH}.

In order to bolster intuition about the relative magnitude of the two effects, we double-sort stocks into deciles based on ownership $z_{Q2}^{\text{RH}}(n)$ and multiplier $\mathcal{M}_{Q2}(n)$. We compute the stock-specific repricing $\Delta_{Q2}^{\text{RH}}(n) = \frac{M_{Q2}(n) - M_{Q2}^{-\text{RH}}(n)}{M_{Q2}(n)}$ for each stock and take averages across stocks within each of the 10×10 Multiplier-Ownership portfolios. Table 3.4 shows the average repricing for all portfolios in percent. In line with the decomposition of Robinhood traders' price impact (3.15), the average repricing conditional on ownership rises from small to big multiplier deciles. Similarly, conditional on the multiplier, the repricing increases with higher ownership. As the overall market share of Robinhood traders is small, these effects are sizeable

[21] In general, a substantial portion of larger stocks' investor base is made up of big institutional investors. Because these investors tend to be fairly inelastic, the micro-multiplier is higher for large-cap stocks. In the Online Appendix we illustrate the distribution of the micro-multiplier across stocks and size deciles.

Table 3.4: **Multiplier-Ownership Double Sorts**
The table reports the average repricing (%) for $10 \times 10 = 10$ multiplier-ownership decile portfolios. As of July 2020, we sort all stocks into Robinhood ownership z_{Q2}^{RH} and multiplier $\mathscr{M}_{Q2}$ deciles. Within each of the 100 portfolios, we compute the average repricing across the N_q stocks in portfolio $q = 1, \ldots, 100$ as $\frac{1}{N_q} \sum_{n \in q} \Delta_{Q2}^{\mathrm{RH}}(n)$, where $\Delta_{Q2}^{\mathrm{RH}}(n) = (M_{Q2}(n) - M_{Q2}^{-\mathrm{RH}}(n))/M_{Q2}(n)$ is the repricing of stock n due to Robinhood traders.

	Ownership Deciles z_{Q2}^{RH}									
Multiplier Deciles $\mathscr{M}_{Q2}$	Low	2	3	4	5	6	7	8	9	High
Small	-0.28	-0.36	-0.33	-0.29	-0.26	0.04	0.54	2.40	9.38	53.67
2	-0.44	-0.45	-0.41	-0.36	-0.28	0.02	0.90	3.32	14.76	59.46
3	-0.46	-0.43	-0.48	-0.36	-0.29	0.04	0.92	3.48	14.04	58.21
4	-0.44	-0.43	-0.40	-0.37	-0.27	0.06	0.99	3.59	15.34	60.50
5	-0.41	-0.45	-0.48	-0.41	-0.23	-0.02	1.03	4.02	13.45	64.70
6	-0.50	-0.49	-0.42	-0.44	-0.21	0.16	1.09	4.45	13.88	53.27
7	-0.45	-0.52	-0.47	-0.39	-0.25	0.13	1.22	4.34	15.84	53.65
8	-0.49	-0.50	-0.51	-0.45	-0.25	0.21	1.62	4.01	20.69	75.54
9	-0.50	-0.54	-0.52	-0.45	-0.22	0.29	1.66	6.23	21.89	76.12
Large	-0.55	-0.57	-0.55	-0.44	-0.22	0.20	2.20	7.11	24.19	89.80

only for the larger ownership deciles. The repricing for the portfolio comprised of the highest Robinhood-ownership stocks that are held by the most inelastic institutions is 94%. This implies that without the presence of Robinhood traders, the average stock in this portfolio would have had a 94% lower market capitalization in Q2.

Lastly, we assess the extraordinary rally of GameStop (GME) during January 2021 through the lens of the demand-based approach. While we do not have data on Robinhood traders' ownership in GameStop during January 2021, we can nevertheless gauge their buying pressure on its share price via the multiplier $\mathscr{M}_t(GME)$. In July 2020, GameStop was already subject to substantial short interest by hedge funds. However, we find that the institutional investors that were long GameStop (as extracted from 13F filings) are primarily inelastic. In fact, over 60% of GameStop's shares outstanding were held by perfectly inelastic (i.e. passive) investors, who do not provide *any* liquidity in the case of collective buying pressure from retail traders.[22] Furthermore, the hedge funds that already had existing short-positions in GameStop were unlikely to provide additional liquidity as this would have increased their short exposure. We estimate a multiplier of $\mathscr{M}_t(GME) = 5.5$, which implies that, as of Q2, buying 10% of Gamestop's shares outstanding causes a 55% increase in GameStop's share price. If GameStop's institutional investor base remained largely unchanged, this multiplier also applies to January 2021.

Thus, the inelastic nature of institutional demand sheds light not only on the impact of the

[22]In the Online Appendix, we plot the price elasticity of demand of GameStop's institutional investor base relative to their ownership shares $z_t^i(GME)$.

retail investment boom on the market crash and recovery of 2020, but also on the extraordinary volatility observed in early 2021.

3.6 Conclusion

This paper investigates the effects of the retail investment boom on the US equity market within a structural model. We find that the majority of all institutional investors - who hold over 60% of the US equity market - have inelastic demand. Because they respond inelastically to price changes, the relatively small retail sector can have a substantial impact on prices. Robinhood traders provided considerable liquidity to the stock market in Q1 and amplified the recovery in Q2. In their absence, the aggregate market capitalization of the smallest quintile of US stocks would have been 25% lower in Q2. While their price impact is concentrated towards smaller stocks, they are able to affect the price of some large companies, which are being held primarily by passive investors. Our findings have important implications for policy makers. Large scale policies, such as the 2020 CARES act under which stimulus checks were sent to most American citizens, have the potential to move prices considerably far from their fundamental values if households invest rather than consume their share. Moreover, the prominent role of Robinhood traders in driving returns evokes concerns about the future role of retail trading in equity markets. If - facilitated by novel fintech solutions - the retail sector continues to grow its wealth share, the extraordinary volatility observed during the pandemic may turn out to be the *new normal*.

3.7 Proofs and Supplementary Material

A Proofs

Proof of Equation (3.4)

Taking the demand function (3.3) and imposing the budget constraint $\sum_n w_t^j(n) = 1 - w_t^j(0)$ yields

$$w_t^j(n) = \frac{\delta_t^j(n)}{1 + \sum_n \delta_t^j(n)}, \tag{3.16}$$

where $\delta_t^j(n) = \exp\{\theta_t^j + \gamma^j me_t(n) + \sum_{k=1}^K \beta_k^j X_{k,t}(n)\}\epsilon_t^j(n)$. Note that $me_t(n) = p_t(n) + s_t(n)$. The vector of investor j's log shares held at t can be written as

$$\mathbf{q}_t^j = \log(A_t^j \mathbf{w}_t^j) - \mathbf{p}_t, \tag{3.17}$$

where bold letters denote vectors in $\mathbb{R}^N$. Differentiating $\mathbf{q}_t^j$ from above with respect to the price vector yields an $N \times N$ matrix[23]

$$\frac{\partial \mathbf{q}_t^j}{\partial \mathbf{p}_t'} = \left(\text{diag}(\mathbf{w}_t^j)\right)^{-1} \frac{\partial \mathbf{w}_t^j}{\partial \mathbf{p}_t'} - \mathbf{I}. \tag{3.18}$$

The diagonal elements of $\frac{\partial \mathbf{w}_t^j}{\partial \mathbf{p}_t'}$ are given by $\frac{\partial w_t^j(n)}{\partial p_t(n)} = \gamma^j w_t^j(n)(1 - w_t^j(n))$, whereas the off-diagonal elements are given by $\frac{\partial w_t^j(n)}{\partial p_t(m)} = -\gamma^j w_t^j(n) w_t^j(m)$. Thus

$$\frac{\partial \mathbf{w}_t^j}{\partial \mathbf{p}_t'} = \gamma^j \left(\text{diag}(\mathbf{w}_t^j) - \mathbf{w}_t^j \mathbf{w}_t^{j\prime}\right). \tag{3.19}$$

Substituting (3.19) into (3.18) yields (3.4).

Proof of Equation (3.15)

Recall the market clearing condition, which implies that $\mathbf{p}_t = \log\left(\sum_j A_t^j \mathbf{w}_t^j(\mathbf{p}_t)\right) - \mathbf{s}_t$. Differentiating both sides with respect to the log latent demand of Robinhood traders $\log(\boldsymbol{\epsilon}_t^{\text{RH}})$ yields

$$\frac{\partial \mathbf{p}_t}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'} = \mathbf{H}_t^{-1} \left(\sum_j A_t^j \frac{\partial \mathbf{w}_t^j}{\partial \mathbf{p}_t'} \frac{\partial \mathbf{p}_t}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'} + A_t^{\text{RH}} \frac{\partial \mathbf{w}_t^j}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'} \right), \tag{3.20}$$

[23] Note that, by assumption, an investor's assets under management A_t^j are determined exogenously and hence do not depend on $\mathbf{p}_t$.

where $\mathbf{H}_t = \sum_j A_t^j \text{diag}(\mathbf{w}_t^j)$. Solving for $\frac{\partial \mathbf{p}_t}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'}$ yields

$$\frac{\partial \mathbf{p}_t}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'} = \left(\mathbf{I} - \sum_j A_t^j \mathbf{H}_t^{-1} \frac{\partial \mathbf{w}_t^j}{\partial \mathbf{p}_t'}\right) A_t^{\text{RH}} \mathbf{H}_t^{-1} \frac{\partial \mathbf{w}_t^{\text{RH}}}{\partial \boldsymbol{\epsilon}_t^{\text{RH}\prime}}, \tag{3.21}$$

where the last term is given by

$$\frac{\partial \mathbf{w}_t^{\text{RH}}}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'} = \text{diag}(\mathbf{w}_t^{\text{RH}}) - \mathbf{w}_t^{\text{RH}} \mathbf{w}_t^{\text{RH}\prime}. \tag{3.22}$$

Substituting (3.19) and (3.22) in (3.21) yields

$$\frac{\partial \mathbf{p}_t}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'} = \left(\mathbf{I} - \sum_j A_t^j \gamma^j \mathbf{H}_t^{-1} \mathbf{G}_t^j\right)^{-1} A_t^{\text{RH}} \mathbf{H}_t^{-1} \mathbf{G}_t^{\text{RH}}, \tag{3.23}$$

where $\mathbf{H}_t = \sum_j A_t^j \text{diag}(\mathbf{w}_t^j)$. The matrix inside the inverse captures the market's demand elasticity, which determines the amplification of demand shocks across the stock universe. If $\text{diag}(\mathbf{w}_j) - \mathbf{w}_j \mathbf{w}_j' \approx \text{diag}(\mathbf{w}_j)$, then $A_t^j \mathbf{H}_t^{-1} \mathbf{G}_t^j = \text{diag}(\mathbf{z}_t^j)$, where $z_t^j(n) = \frac{w_t^j(n) A_t^j}{\sum_k w_t^k(n) A_t^k}$ is the ownership share of j in stock n. Thus, we can write

$$\frac{\partial \mathbf{p}_t}{\partial \log(\boldsymbol{\epsilon}_t^{\text{RH}})'} = \left(\mathbf{I} - \sum_j \gamma^j \text{diag}(\mathbf{z}_t^j)\right)^{-1} \text{diag}(\mathbf{z}_t^{\text{RH}}). \tag{3.24}$$

B Institutional Holdings

Institutional Investors

We obtain quarterly institutional holdings data on stocks traded on the New York Stock Exchange (NYSE), the NYSE American, and the NASDAQ between January 2005 and July 2020 from Refinitiv. We restrict our sample to holdings disclosed in 13F filings. Note that Refinitiv Ownership reports holdings by investment managers. An investment manager does not always correspond to an entire institution, but can be a subsidiary.[24] Using Refinitiv, we identify the parent institution of each investment manager and aggregate holdings at the parent institution level. We compute the percentage of ownership as the number of shares held divided by shares outstanding. Following Ben-David et al. (2016), we winsorize ownership shares above 50% and proportionally scale down holdings so that total institutional ownership does not exceed 100%.[25] We construct each institution's equity portfolio weights as the dollar holdings in each stock (price times shares held) divided by their assets under management. We compute assets under management as the sum of an institution's dollar holdings. Following KY (2019), we

[24] For example, BlackRock is split between 16 investment managers.

[25] Lewellen (2011) finds that institutional holdings exceeding 100% of shares outstanding are not data errors but the result of double-counting shares that were short-sold.

Table 3.1: **Institutional Ownership Data Summary**
This table reports summary statistics on institutional holdings. The number of financial institutions (first column) is computed after the cleaning steps *A*, *B*, and *C*. This number is the count of institutions present in our sample in either or both Q1 and Q2. The remaining statistics displayed in the last 7 columns are the average of the Q1 and Q2 quarterly statistics.

	Number of FI	Percentage of Market Held	Assets under Management		Number of Stocks Held		Number of Stocks in Universe	
			Median	90th Percentile	Median	90th Percentile	Median	90th Percentile
Banks	170	7	599	37,016	187	1,387	319	1,896
Insurance Companies	34	2	703	10,309	109	591	200	860
Pension Funds	62	3	6,434	42,153	704	1,768	971	2,219
Investment Advisors	2,822	50	194	3,097	82	439	147	833
Hedge Funds	512	2	315	2,458	28	211	79	573
Other	50	1	677	15,812	50	1,517	106	2,367

attribute the holdings of institutions with less than $10 million under management to the household sector. Similarly, all institutions that do not have any holdings in any inside assets or the outside asset are merged with households.

We regroup Refinitiv's 14 investor categories into six investor types. Insurance companies and hedge funds remain as they are. Banks and trusts are grouped with holding companies, investment advisors with investment advisors/hedge funds, and pension funds with sovereign wealth funds. Finally, endowment funds, venture capital firms, research firms, foundations, private equity firms, and corporations are grouped together (labelled as *Other*). Table 3.1 reports summary statistics on institutional holdings by investor type for the first half of 2020.

Sparse Portfolios

Estimating the demand system requires estimating demand coefficients for each individual investor. However, most of the institutions in our sample hold very concentrated portfolios or have filed few 13F filings. Indeed, the majority of institution-quarter pairs report less than 100 holdings. In order to obtain sufficiently many observations for the demand estimation, we pool institutions with insufficient holdings together and estimate demand coefficients for the pooled entities. We establish two cutoffs; i) a cross-sectional cutoff of 100 positive holdings in our investment universe (excluding the outside asset) that must be satisfied in every quarter, and ii) a time series cutoff of 12 quarters of available 13F filings. Institutions that do not meet the required cutoffs are only pooled together with institutions of the same investor type. Additionally, investment advisors are pooled with investment advisors that share the same investment style. We obtain advisors' investment style from Refinitiv. Within each type (banks, hedge funds, insurance companies, pension funds, and other) and investment style (core value, core growth, growth at a reasonable price (GARP), growth, hedge fund, deep value, income value, yield, and other), we sort institutions by quantiles of average AUM. The holdings of the institutions belonging to the same type/style and quantile are summed together and form the aggregated holdings of our pooled entities. The number of quantiles is set such that the cross-sectional and the time series cutoffs are satisfied for every pool. Note that, to conform with our panel approach, pools are immutable. Thus, an institution is uniquely assigned to the same pool in every quarter.

C Robinhood Holdings

Summary Statistics

Robintrack reports $H_t(n)$, i.e. the number of accounts holding each stock traded on the online brokerage platform "Robinhood Markets Inc.". Table 3.1 reports summary statistics on $H_t(n)$ for Q1 and Q2.

Table 3.1: **Robinhood Accounts Holding Summary Statistics**
This table reports summary statistics on the number of Robinhood accounts holding a stock $H_t(n)$ for the last quarter of 2019, and the first two of 2020. The summary statistics are cross-sectional measures computed across all stock within our investment universe. For example, the mean corresponds to $\frac{1}{N}\sum_{n=1}^{N} H_t(n)$. The number of observations corresponds to the number of stocks traded on Robinhood. Changes in $H_t(n)$ are computed for each stock-quarter pair in Q1 and Q2.

	Mean	Std	Min	P25	Median	P75	Max	N(Obs.)
Accounts Holding $H_t(n)$								
2019 Q4	3,070	17,168	0	118	364	1,196	321,191	1,981
2020 Q1	5,667	30,147	0	173	543	1,916	649,148	2,028
2020 Q2	9,601	48,158	0	276	906	3,158	925,539	2,076
Changes in Accounts Holding $H_t(n) - H_{t-1}(n)$								
2020 Q1	2,602	15,806	-4,394	17	93	507	327,957	1,903
2020 Q2	3,929	21,304	-11,138	50	213	922	397,892	2,016

Portfolio Weight Approximation

Note that $H_t(n)$ cannot be readily interpreted as the total number of shares held by all Robinhood traders in the respective stock. This is because $H_t(n)$ simply tracks the number of registered accounts holding stock n, regardless of whether an account holds multiple shares or a fraction of a share. However, we can use the number of accounts holding each stock to approximate the representative Robinhood portfolio as in (3.1). To this end, let $W_{a,t}$ denote the wealth of account a on Robinhood and $w_{a,t}(n)$ the account's portfolio weight in stock n.

Then, the aggregate Robinhood portfolio is given by

$$w_t^{\mathrm{RH}}(n) = \frac{\sum_{a=1}^{U_t} W_{a,t} w_{a,t}(n)}{\sum_{a=1}^{U_t} W_{a,t}}. \tag{3.25}$$

If every account holding a stock $\mathbb{1}_{a,n}(n)$ represents an equal amount of dollars USD_t, then

$$\frac{\sum_{a=1}^{U_t} W_{a,t} w_{a,t}(n)}{\sum_{a=1}^{U_t} W_{a,t}} = \frac{\sum_{a=1}^{U_t} \mathbb{1}_{a,n}(n) USD_t}{\sum_{a=1}^{U_t} \sum_{n=1}^{N} \mathbb{1}_{a,t}(n) USD_t} = \frac{H_t(n)}{\sum_{n=1}^{N} H_t(n)}. \tag{3.26}$$

The approximation also holds if $\mathbb{1}_{a,t}(n) = 1$ represents different dollar amounts across users, as long as the wealth distribution is uncorrelated with the users' investment universes, i.e. their sets of preferred stocks. Note that, because users can buy fractional shares on Robinhood, the actual dollar price of a stock plays a subordinate role for the investment decision, even under a limited budget. Lastly, the approximation even holds if $\mathbb{1}_{a,t}(n) = 1$ represents different dollar amounts across stocks, as long as investors' deviations from holding equal-weighted portfolios are random. If a user randomly distributes her budget across stocks within her investment universe, then (as U_t grows large) the average user holds an equal-weighted portfolio within her universe and (3.1) holds.

Scaling the Implied Holdings

While the actual amount of shares held by Robinhood traders $Q_t^{\mathrm{RH}}(n)$ is unobserved, we reverse-engineer an approximation using (3.1) and the AUM figures of the S-1 filing from July 1st, 2021.

$$Q_t^{\mathrm{RH}}(n) = \frac{w_t^{\mathrm{RH}}(n) A_t^{\mathrm{RH}}}{P_t(n)}. \tag{3.27}$$

The S-1 filing only reports key metrics for December 31st, 2019, as well as March 31st and December 31st, 2020. In order to infer the shares held in Q2, approximate need to Robinhood's AUM given the reported figures in the S-1. The publicly reported user base of 13 Million in May 2020 is roughly equivalent to the number of disclosed users in the S-1 for December 31st, 2020. While the rising number of open positions (Figure 3.1) suggests that Robinhood traders' AUM likely increased from May to July 2020, we take a conservative approach and assume their AUM to be \$50 billion throughout Q2. In order to get AUM approximations for all quarters before December 2019, that are consistent the evolution of open positions, we fit a polynomial to $\frac{1}{N}\sum_{n=1}^{N} H_t(n)$ and evaluate the function at different dates, given that in December 31st it took \$11.7 billion. The residual shares not held by Robinhood traders or institutions make up the household sector. In the rare cases, where institutional and implied Robinhood holdings exceed the shares outstanding, we rescale Robinhood traders' shares held to be $\tilde{Q}_t^{\mathrm{RH}}(n) = S_t - \sum_{i=1}^{I} Q_t^i(n)$. Robinhood portfolio weights are then recomputed as $w_t^{\mathrm{RH}}(n) = \frac{\tilde{Q}_t^{\mathrm{RH}}(n) P_t(n)}{\sum_{n=1}^{N} \tilde{Q}_t^{\mathrm{RH}}(n) P_t(n)}$.

D Quarterly Demand Estimation

KY (2019) propose the following demand function, which is purely cross-sectional and implies time-varying demand coefficients,

$$\forall j,t: \quad \frac{w_t^j(n)}{w_t^j(0)} = \exp\left\{\beta_{0,t} + \gamma_t^j me_t(n) + \sum_{k=1}^{K} \beta_{k,t}^j X_{k,t}(n)\right\}\epsilon_t^j(n). \tag{3.28}$$

Thus, each investor's demand is estimated over the cross-section of portfolio weights at time t and is re-estimated quarterly. In our setting, this implies the following quarterly moment conditions for 13F-institutions, households and Robinhood traders,

$$\forall j,t: \quad \mathrm{E}\left[\log(\epsilon_t^j(n)) \,\middle|\, \widehat{me}_t^j(n), X_t(n)\right] = 0 \quad s.t. \quad \gamma_t^j < 1, \tag{3.29}$$

from which we estimate demand by constrained generalized method of moments. Note, that for each investor, we obtain a time series of coefficients. Robust inference from counterfactual experiments (e.g. redistributing Robinhood's assets under management) requires that the estimated coefficients do not change in the counterfactual. If the estimated coefficients vary strongly over time, it implies demand curves are inherently unstable as they tend to vary with the macro-economic or political environment. Arguing for counterfactual-invariant demand coefficients then becomes difficult. Having estimated $\{\gamma_t^j\}_{t=1}^T$ and $\{\beta_{k,t}^j\}_{t=1}^T$ for all characteristics k and for all quarters $t = 1,...,T$, we compute the wealth-weighted average across investors within each investor type at each quarter. Using the time series of wealth-weighted coefficients for each type, we compute the mean and standard deviation over time. Table 3.1 reports the estimated coefficients and their time series standard deviations for each investor type.

While the average coefficients are roughly equivalent to the panel estimates, many are statistically insignificant. Note that the averages of quarterly re-estimated coefficients are not necessarily equivalent to the reported panel estimates due to the coefficient constraint ($\gamma^j < 1$). Furthermore, coefficient averages are wealth-weighted at all dates, whereas Table 3.1 reports coefficient averages using wealth shares as of Q2 only.

Table 3.1: **Quarterly Demand Coefficients**
The table reports the average coefficients and their time series standard deviations for the quarterly re-estimated demand curves. Within each investor type, we first compute AUM-weighted average coefficients for each quarter and then compute time series averages and standard deviations across quarters.

		Demand Coefficients					
		γ^j	β^j_{be}	$\beta^j_{\text{inv.}}$	β^j_{profit}	β^j_{beta}	$\beta^j_{\text{div/be}}$
Institutional Investors							
Hedge Funds	Mean	0.53	-4.64	9.12	4.45	7.39	-6.26
	Std.	0.07	3.18	2.73	7.45	5.06	7.01
Insurance Companies	Mean	0.72	0.55	4.26	-5.09	-8.23	2.60
	Std.	0.06	4.73	5.73	14.68	7.43	13.81
Other	Mean	0.82	22.57	4.70	4.10	0.61	-0.43
	Std.	0.10	13.66	4.65	7.13	5.72	6.93
Investment Advisors	Mean	0.82	13.67	3.11	4.84	-0.09	-2.92
	Std.	0.03	1.74	2.11	7.32	2.63	6.06
Banks	Mean	0.89	16.70	0.57	2.50	-6.29	-0.85
	Std.	0.04	4.93	2.83	6.24	4.76	5.28
Pension Funds	Mean	0.94	5.98	-0.93	3.81	-1.23	-2.73
	Std.	0.01	2.19	1.91	10.77	2.57	10.52
Households	Mean	0.72	9.50	-1.11	-0.14	-11.86	6.09
	Std.	0.03	4.47	2.84	6.42	4.44	6.28
Robinhood Traders	Mean	0.32	-27.07	7.32	-7.41	24.25	8.29
	Std.	0.03	1.69	4.27	6.91	16.33	5.48

Conclusion

This thesis studies the incorporation of data on investors' portfolio holdings into equilibrium asset pricing models. The three chapters investigate the theory, estimation and applications of such demand-based models. In particular, I study the application and structural estimation of demand systems in asset pricing. A series of papers studies 1) the identification of demand curves from data on portfolio holdings, 2) the implications of institutional demand for the pricing of sustainable investments, and 3) the interaction of retail traders and large institutions.

There is a wealth of data on investor-level portfolio holdings and capital flows across assets, asset classes and countries. If demand curves for financial securities are downward sloping, then one can learn a lot about asset prices by linking them to capital flows and portfolio holdings through equilibrium models. In this light, the use of quantity data provides an explicit link between market micro-structure, intermediary asset pricing, and corporate finance. The market micro-structure literature has carefully developed theories and tools to estimate downward-sloping demand curves, price impact, and the relationship between prices and flows at a higher frequency. Investigating how investors' demand shocks enter prices at a higher frequency and estimating their permanent impact is an exciting avenue for future research. Another interesting direction for future research is understanding the interaction of investor demand and the supply curve, i.e. the decisions at the corporate level. What is the joint role of investor demand, downward sloping demand curves, capital structure and investment decisions? An explicit link between asset pricing and corporate finance via the use of portfolio holdings would allow quantifying the real effects of of financial markets. Do sustainable firms capitalize on the rise of sustainable investing by issuing new shares at elevated prices and investing in green projects?

Overall, demand-based asset pricing opens up an exciting research agenda which, on the one hand, develops the tools to incorporate quantity data into equilibrium asset pricing models, and on the other hand, uses these tools to uncover new facts about the sources of financial market fluctuations. Understanding the role of quantities and prices jointly will provide valuable insights for policymakers as it allows making quantitative statements about the level of intervention required to affect asset prices in the desired direction.

Bibliography

Balasubramaniam, V., J. Y. Campbell, T. Ramadorai, and B. Ranish (2021). Who owns what? a factor model for direct stock holding. Technical report, National Bureau of Economic Research.

Baldauf, M., L. Garlappi, and C. Yannelis (2020). Does climate change affect real estate prices? only if you believe in it. *The Review of Financial Studies 33*(3), 1256–1295.

Bams, D. and B. van der Kroft (2022). Divestment, information asymmetries, and inflated esg ratings. *Available at SSRN 4126986*.

Barber, B. M., X. Huang, T. Odean, and C. Schwarz (2020). Attention induced trading and returns: Evidence from Robinhood users. Working paper.

Barber, B. M., T. Odean, and N. Zhu (2008). Do retail trades move markets? *The Review of Financial Studies 22*(1), 151–186.

Barrot, J.-N., R. Kaniel, and D. Sraer (2016). Are retail traders compensated for providing liquidity? *Journal of Financial Economics 120*(1), 146–168.

Ben-David, I., F. Franzoni, R. Moussawi, and J. Sedunov (2016, May). The granular nature of large institutional investors. Working Paper 22247, National Bureau of Economic Research.

Ben-David, I., J. Li, A. Rossi, and Y. Song (2020). Advice-driven demand and systematic price fluctuations. Technical report, National Bureau of Economic Research.

Benetton, M. and G. Compiani (2020). Investors' beliefs and asset prices: A structural model of cryptocurrency demand. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-107).

Berg, F., J. F. Koelbel, and R. Rigobon (2019). *Aggregate confusion: The divergence of ESG ratings.* MIT Sloan School of Management.

Berg, F., J. F. Kölbel, A. Pavlova, and R. Rigobon (2021). Esg confusion and stock returns: Tackling the problem of noise. *Available at SSRN 3941514*.

Berk, J. and J. H. van Binsbergen (2022). The impact of impact investing. *Available at SSRN 3909166*.

Bernstein, A., M. T. Gustafson, and R. Lewis (2019). Disaster on the horizon: The price effect of sea level rise. *Journal of financial economics 134*(2), 253–272.

Betermier, S., L. E. Calvet, S. Knüpfer, and J. Kvaerner (2022). What do the portfolios of individual investors reveal about the cross-section of equity returns? *Available at SSRN 3795690*.

Blume, M. E., D. B. Keim, et al. (2014). The changing nature of institutional stock investing. *Critical Finance Review 6*, 1–41.

Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang (2020). Tracking retail investor activity. *Journal of Finance forthcoming*.

Bolton, P. and M. Kacperczyk (2021a). Do investors care about carbon risk? *Journal of Financial Economics*.

Bolton, P. and M. Kacperczyk (2021b). Global pricing of carbon-transition risk. Technical report, National Bureau of Economic Research.

Bretscher, L., L. Schmid, I. Sen, and V. Sharma (2020). Institutional corporate bond pricing. *Swiss Finance Institute Research Paper* (21-07).

Chang, Y.-C., H. Hong, and I. Liskovich (2015). Regression discontinuity and the price effects of stock market indexing. *The Review of Financial Studies 28*(1), 212–246.

Chen, H. (2020). Cash induced demand. *Unpublished Working Paper*, 62.

Coval, J. and E. Stafford (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics 86*(2), 479–512.

Cremers, K. M. and A. Petajisto (2009). How active is your fund manager? a new measure that predicts performance. *The review of financial studies 22*(9), 3329–3365.

Derrien, F., P. Krueger, A. Landier, and T. Yao (2021). How do esg incidents affect firm value? *Available at SSRN 3903274*.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of finance 65*(4), 1237–1267.

Eaton, G. W., T. C. Green, B. Roseman, and Y. Wu (2021). Zero-commission individual investors, high frequency traders, and stock market quality. Working paper.

Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial economics 101*(3), 621–640.

Edmans, A., I. Goldstein, and W. Jiang (2012). The real effects of financial markets: The impact of prices on takeovers. *The Journal of Finance 67*(3), 933–971.

Faccini, R., R. Matin, and G. S. Skiadopoulos (2021). Dissecting climate risks: Are they reflected in stock prices? *Available at SSRN 3795964*.

Fama, E. F. and K. R. French (2007). Disagreement, tastes, and asset prices. *Journal of financial economics 83*(3), 667–689.

Foucault, T., D. Sraer, and D. J. Thesmar (2011). Individual investors and volatility. *The Journal of Finance 66*(4), 1369–1406.

Frazzini, A. and O. A. Lamont (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of financial economics 88*(2), 299–322.

Gabaix, X. and R. S. Koijen (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. *Available at SSRN 3686935*.

Gibson, R., S. Glossner, P. Krueger, P. Matos, and T. Steffen (2022). Do responsible investors invest responsibly? *Swiss Finance Institute Research Paper* (20-13).

Glossner, S. (2021). Repeat offenders: Esg incident recidivism and investor underreaction. *Available at SSRN 3004689*.

Glossner, S., P. Matos, S. Ramelli, and A. F. Wagner (2020). Where do institutional investors seek shelter when disaster strikes? Evidence from COVID-19. Working Paper DP15070.

Goldsmith-Pinkham, P. S., M. Gustafson, R. Lewis, and M. Schwert (2021). Sea level rise exposure and municipal bond yields. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Görgen, M., A. Jacob, M. Nerlinger, R. Riordan, M. Rohleder, and M. Wilkens (2020). Carbon risk. *Available at SSRN 2930897*.

Greenwood, R. (2005). Short-and long-term demand curves for stocks: theory and evidence on the dynamics of arbitrage. *Journal of Financial Economics 75*(3), 607–649.

Greenwood, R. and D. Thesmar (2011). Stock price fragility. *Journal of Financial Economics 102*(3), 471–490.

Grossman, S. J. and M. H. Miller (1988). Liquidity and market structure. *the Journal of Finance 43*(3), 617–633.

Haddad, V., P. Huebner, and E. Loualiche (2021). How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing. *Theory, Evidence from Portfolios, and Implications for the Rise of Passive Investing (April 7, 2021)*.

Han, X., N. L. Roussanov, and H. Ruan (2021). Mutual fund risk shifting and risk anomalies. *Available at SSRN 3931449*.

Hartzmark, S. M. and D. H. Solomon (2021). Predictable price pressure. *Available at SSRN 3853096*.

Hong, H. and M. Kacperczyk (2009). The price of sin: The effects of social norms on markets. *Journal of financial economics 93*(1), 15–36.

Hong, H., F. W. Li, and J. Xu (2019). Climate risks and market efficiency. *Journal of econometrics 208*(1), 265–281.

Hsu, P.-H., K. Li, and C.-Y. Tsou (2020). The pollution premium. *Available at SSRN 3578215.*

Huynh, T. D. and Y. Xia (2021). Climate change news risk and corporate bond returns. *Journal of Financial and Quantitative Analysis 56*(6), 1985–2009.

Ilhan, E., Z. Sautner, and G. Vilkov (2021). Carbon tail risk. *The Review of Financial Studies 34*(3), 1540–1571.

In, S. Y., K. Y. Park, and A. Monk (2020). Is 'being green'rewarded in the market? an empirical investigation of decarbonization risk and stock returns. *International Association for Energy Economics (Singapore Issue) 46*, 48.

Jansen, K. A. (2021). Long-term investors, demand shifts, and yields. *Available at SSRN 3901466.*

Jiang, Z., R. Richmond, and T. Zhang (2020). A portfolio approach to global imbalances. *NYU Stern School of Business.*

Kaniel, R., G. Saar, and S. Titman (2008). Individual investor trading and stock returns. *The Journal of finance 63*(1), 273–310.

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

Kim, S. and A. Yoon (2022). Analyzing active fund managers' commitment to esg: Evidence from the united nations principles for responsible investment. *Management Science.*

Koijen, R. S., R. J. Richmond, and M. Yogo (2022). Which investors matter for global equity valuations and expected returns? *Available at SSRN*.

Koijen, R. S. and M. Yogo (2019). A demand system approach to asset pricing. *Journal of Political Economy 127*(4), 1475–1515.

Koijen, R. S. and M. Yogo (2020). Exchange rates and asset prices in a global demand system. Technical report, National Bureau of Economic Research.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315–1335.

Lettau, M., S. C. Ludvigson, and P. Manoel (2018). Characteristics of mutual fund portfolios: Where are the value funds? Technical report, National Bureau of Economic Research.

Lewellen, J. (2011). Institutional investors and the limits of arbitrage. *Journal of Financial Economics 102*(1), 62–80.

Liang, H., L. Sun, and M. Teo (2021). Greenwashing: Evidence from hedge funds.

Lou, D. (2012). A flow-based explanation for return predictability. *The Review of Financial Studies 25*(12), 3457–3489.

Noh, D. and S. Oh (2020). Are green investors green-inducing? a demand system approach. *A Demand System Approach (June 30, 2020).*

Noh, D. and S. Oh (2022). Measuring institutional pressure for greenness: A demand system approach. *Available at SSRN 3639693.*

Ozik, G., R. Sadka, and S. Shen (2020). Flattening the illiquidity curve: Retail trading during the COVID-19 lockdown. Working paper.

Pagano, M. S., J. Sedunov, and R. Velthuis (2020). How did retail investors respond to the COVID-19 pandemic? The effect of Robinhood brokerage customers on market quality. Working paper.

Painter, M. (2020). An inconvenient cost: The effects of climate change on municipal bonds. *Journal of Financial Economics 135*(2), 468–482.

Parker, J. A., A. Schoar, and Y. Sun (2020). Retail financial innovation and stock market dynamics: The case of target date funds. Technical report, National Bureau of Economic Research.

Pástor, L., R. F. Stambaugh, and L. A. Taylor (2021). Sustainable investing in equilibrium. *Journal of Financial Economics.*

Pástor, L., R. F. Stambaugh, and L. A. Taylor (2022). Dissecting green returns. *Journal of Financial Economics 146*(2), 403–424.

Pavlova, A. and T. Sikorskaya (2022). Benchmarking intensity.

Pedersen, L. H., S. Fitzgibbons, and L. Pomorski (2020). Responsible investing: The esg-efficient frontier. *Journal of Financial Economics.*

Peress, J. and D. Schmidt (2020). Glued to the tv: Distracted noise traders and stock market liquidity. *The Journal of Finance 75*(2), 1083–1133.

Petajisto, A. (2009). Why do demand curves for stocks slope down? *Journal of Financial and Quantitative Analysis 44*(5), 1013–1044.

Schmickler, S. (2020). Identifying the price impact of fire sales using high-frequency surprise mutual fund flows. *Available at SSRN 3488791.*

Schmickler, S. N. and P. Tremacoldi-Rossi (2022). Spillover effects of payouts on asset prices and real investment.

Seltzer, L. H., L. Starks, and Q. Zhu (2022). Climate regulatory risk and corporate bonds. Technical report, National Bureau of Economic Research.

Shleifer, A. (1986). Do demand curves for stocks slope down? *The Journal of Finance 41*(3), 579–590.

Stock, J. H. and M. Yogo (2002). Testing for weak instruments in linear IV regression. Working paper, National Bureau of Economic Research.

van der Beck, P. (2022). Identifying elasticities in demand-based asset pricing. *Available at SSRN*.

van der Beck, P. and C. Jaunin (2021). The equity market implications of the retail investment boom. *Swiss Finance Institute Research Paper* (21-12).

Wardlaw, M. (2020). Measuring mutual fund flow pressure as shock to stock returns. *The Journal of Finance 75*(6), 3221–3243.

Welch, I. (2021). The wisdom of the Robinhood crowd. *Journal of Finance forthcoming*.

Wurgler, J. and E. Zhuravskaya (2002). Does arbitrage flatten demand curves for stocks? *The Journal of Business 75*(4), 583–608.

Yang, R. (2021). What do we learn from ratings about corporate social responsibility (csr)? *Columbia Business School Research Paper* (18-37).

PHILIPPE VAN DER BECK

Lausanne, Switzerland | +41 (0)766048573 | philippe.vanderbeck@epfl.ch | pvanderbeck.net

EDUCATION

École Polytechnique Fédérale de Lausanne *2018 - 2023*
PhD in Finance - Joint Program with the Swiss Finance Institute

Princeton University *Fall 2022*
Visiting Student - Department of Economics & Bendheim Center for Finance

Imperial College Business School *2017 - 2018*
MSc Finance - (Top of the class 2018)

Ludwig Maximilians University Munich *2014 - 2017*
BSc Economics - (Top of the class 2017)

WORKING PAPERS

1. *"Flow-Driven ESG Returns"* (JMP, solo-authored)
    - Will be presented at the AFA 2023: *Sustainable Finance and Asset Prices*
    - Winner of the Swiss Finance Institute Best Paper Award 2023
2. *"On the Estimation of Demand-Based Asset Pricing Models"* (solo-authored)
3. *"The Equity Market Implications of the Retail Investment Boom"* (with Coralie Jaunin)
    - Winner of the Swiss Finance Institute Best Paper Award 2022
4. *The Equilibrium Flow-Return Relation* (with Semyon Malamud and Andreas Schrimpf)
5. *Portfolio Holdings and the Origins of Demand Elasticities* (with Lorenzo Bretscher)

AWARDS AND HONORS

Best Paper Award 2022 SFI Research Days *2022*
- Paper: *Flow-Driven ESG Returns*

Best Paper Award 2021 SFI Research Days *2021*
- Paper: *The Equity Market Implications of the Retail Investment Boom*

Best Discussant Award 2020 SFI Research Days *2020*

Best Research Project Prize Imperial College Msc Finance *2019*
- Paper: *High Frequency Decomposition of the Equity Risk Premium*

Award for Highest GPA in MSc Finance *2019*
- Imperial College Business School

Nigel Meade Quantitative Finance Prize *2019*

Unigestion Investment Prize *2019*

Alumni Award for Young Economists LMU Munich *2017*
- Best GPA in BSc Economics of 2017

GRANTS AND SCHOLARSHIPS

Participation in SNSF Grant obtained by Pierre Collin-Dufresne *2020-2023*
- Project: Demand-based Asset Pricing and ESG Preferences
- Amount: 475,490CHF

Imperial College Brilliant Minds Scholarship *2017-2018*
- *Amount: 20,000 GBP*

Swiss Finance Institute *2018 - 2023*
- *PhD Student Scholarship*

Max Weber Program Bavaria *2015 - 2018*
- *Undergraduate Elite Student Scholarship*

PROFESSIONAL ACTIVITIES

Capital Fund Management (CFM, Paris) *2021-2023*
- Academic Consultant for Demand-Based Asset Pricing

Max Planck Institute for Innovation and Competition (MPI, Munich) *2015-2017*
- *Research Assistant*

Ifo Center for Labor and Demographic Economics (CESifo, Munich) *2016-2017*
- *Research Assistant*

REFEREEING

Financial Analysts Journal (FAJ)
Journal of International Financial Markets, Institutions & Money
Economic Modelling

PRESENTATIONS

American Finance Association Meeting *2023*

Northern Finance Association Meeting, Princeton PhD Workshop, World Finance Conference Turin, Dauphine Finance PhD Workshop, HEC Lausanne PhD Workshop, SFI Research Days *2022*

14th Financial Risk International Forum, Society of Quantitative Analysts (SQA), SFI Research Days, Unil Brownbag Seminar, Wharton PhD Workshop (co-author presentation), CFM Research Seminar (x2), Vrije University Amsterdam Finance Seminar, SFI Research Days *2021*

SFI Research Days *2020*

TEACHING ASSISTANTSHIPS

Real Options and Financial Structuring, MSc Financial Engineering, EPFL *2019-2023*
Risk Management Using Factor Models, SFI Master Class *2021-2022*
Data-Driven Business Analytics, Humanities and Social Sciences Program, EPFL *2021-2022*
Advanced Microeconomics, BSc Economics, LMU *2016-2017*

SKILLS & INTERESTS

Software: Python, Stata, Matlab, LaTeX
Statistics & Machine Learning: Causal Inference, Structural Estimation, Deep Learning
Interests: Tennis, Track & Field, CrossFit, Golf

REFERENCES

Pierre Collin-Dufresne
Professor of Finance
EPFL
Swiss Finance Institute
+41 (0)21 693 0136
pierre.collin-dufresne@epfl.ch

Motohiro Yogo
Professor of Economics
Princeton University
Bendheim Center for Finance
+1 609 258 4467
myogo@princeton.edu

Lorenzo Bretscher
Assistant Professor of Finance
HEC Lausanne
Swiss Finance Institute
+41 (0)21 692 61 29
lorenzo.bretscher@unil.ch