The Impact Frontier

PRELIMINARY DRAFT

Please do not cite or circulate without permission

JONATHAN HARRIS*

January 31, 2026

ABSTRACT

Asset pricing usually takes firms' demand for capital as given, while corporate finance does the same for investors' supply of capital. In a unified model, what real effects can investor portfolio tilts have on climate- and nature-related externalities over multi-year horizons? This paper develops such an equilibrium model and constructs the Impact Frontier: the Pareto frontier between certainty-equivalent returns and the expected effect on an externality. Portfolios on the impact frontier are obtained by allocating to an impact-tilt portfolio defined by additionality-adjusted externality scores—raw scores (illustrated by ESG ratings) multiplied by a matrix derived from demand- and supply-side elasticities. Calibrated to U.S. equity data, the impact frontier differs materially from the ESG-efficient frontier; portfolios based only on raw scores can lie far below the frontier and even yield negative expected impact.

*MIT Sloan School of Management and the Total Portfolio Project at the Center for Sustainable Finance and Private Wealth, Zurich, Switzerland; jonathan@total-portfolio.org. I am grateful for helpful comments and suggestions on the development of this paper from Sebastien Betermier, Robert C. Merton, Laurent Calvet, Marcin Kacperczyk, Motohiro Yogo, Philippe van der Beck, Raman Uppal, Florian Heeb, Jules van Binsbergen, Florian Berg, Bram van der Kroft, Jon Lukomnik, Jim Hawley, Costanza Consolandi, Paul Huebner, Olivier David Zerbib, Laurent Barras, Jan Starmans, Julian Kölbel, Roberto Rigobon, Don Noh, Lea Tschan, Falko Paetzold, John Firth, Ellen Quigley, Riccardo Rebonato, Mike McCreless, Harald Walkate, Paddy Carter, Paul Christiano, Wayne Chang, Eric Tham, John Berger, and participants at the 3rd Spring Finance Workshop, the Conference on Climate and Energy Finance at Leibniz University Hannover, and the 2023 GRASFI Annual Conference at the Yale SOM.

Acknowledgements and disclosures. I gratefully acknowledge financial support for this research from Stichtung ENDEE, Jazi Foundation, the CSP Foundation, and the Externality Investment Research Network. I have no financial relationships or other potential conflicts of interest to disclose.

1 Introduction

A central question in climate and nature finance is whether, and through which mechanisms, investor portfolio choices can have real effects on firms' activities and the associated externalities. A common modeling divide is that asset pricing often takes firms' demand for capital as given, while corporate finance takes investors' supply of capital as given.[1] This paper develops a unified equilibrium model of these two sides and uses it to construct the *Impact Frontier*: the Pareto frontier between certainty-equivalent returns and the expected effect on an externality.

The paper focuses on the capital-allocation channel of portfolio tilts and how they transmit into installed capital and an externality; related channels such as shareholder engagement and governance are complementary mechanisms but are not the focus of the model. In this framework, "investor impact" is defined as the expected change in the equilibrium level of an aggregate externality caused by a marginal shift in an investor's supply schedule (Brest, Gilson, and Wolfson, 2018; Kölbel et al., 2020). A key challenge is that the same portfolio tilt can have very different real effects across firms because the pass-through from investor supply of capital to firm scale depends on both the slope of investor supply and the slope of firms' demand for capital.

The paper's first contribution is to provide a tractable microfounded supply-and-demand system in which these slopes (and hence real effects) are allowed to vary in the cross section. The central result is that portfolios designed to maximize investor impact for a given certainty-equivalent return objective will generally differ from portfolios that tilt only on *raw externality scores*—for example, firm-level ESG ratings. In particular, ignoring heterogeneity in pass-through can move an investor away from the impact frontier and, in calibrated examples, can even imply negative investor impact despite tilting toward firms with higher raw scores.

The model builds on the supply-and-demand approach in Betermier, Calvet, and Jo (2025,

[1] Following Betermier, Calvet, and Jo (2025), I use a "capital" orientation: investors supply capital and firms demand capital. An equivalent "asset" convention frames investors as the demand side and firms as the supply side. This is a notational convention and does not change results.

hereafter BCJ), Merton (1987); Zhang (2017); Koijen and Yogo (2019); Pástor, Stambaugh, and Taylor (2021); Van der Beck (2022), and responds to calls for theory that accommodates non-traditional investor preferences in sustainable finance (Starks, 2023). Relative to BCJ, the model adds (i) an explicit externality that depends on equilibrium firm scale, and (ii) cross-sectional heterogeneity in supply- and demand-side elasticities that governs the strength of real-economy responses. The impact frontier then provides a compact way to summarize optimal portfolio policies for investors with different "prices of impact" (i.e., marginal rates of substitution between the externality and wealth).

In a single-firm version of the model, the elasticity of firm capital with respect to a marginal investor supply shift is summarized by a *contribution multiplier* $\mathcal{C}$,

$$\mathcal{C} = \frac{\delta K}{\delta K_i} \approx \frac{\mathcal{M}_S}{\mathcal{M}_S + \mathcal{M}_D}, \tag{1}$$

where $\mathcal{M}_S$ is the slope of the aggregate investor supply curve and $\mathcal{M}_D$ is the slope of the firm's demand curve, as illustrated in Figure 1. The multiplier is near zero in endowment-style settings with effectively inelastic firm demand ($\mathcal{M}_D \to \infty$) and near one in settings with highly elastic firm demand ($\mathcal{M}_D \to 0$). Allowing $\mathcal{M}_S$ and $\mathcal{M}_D$ to vary across firms yields economically meaningful heterogeneity in $\mathcal{C}$ and, consequently, in investor impact.

In the full model, equilibrium spillovers imply a matrix of contribution multipliers, so that a supply shift toward one firm can affect the equilibrium scale of other firms. Portfolios on the impact frontier are obtained by allocating to an *impact-tilt portfolio* defined by *additionality-adjusted externality scores*—raw scores multiplied by the contribution multiplier matrix.

The framework also clarifies the relationship between "holdings-based" sustainable-investing motives and consequentialist, impact-based motives (Yasuda, 2023). Holdings-based preferences depend only on the investor's own holdings and can be represented as exogenous tilts. In contrast, impact preferences in this paper are defined over an aggregate externality that depends on equilibrium firm scale, so optimal portfolio choice depends on the supply-and-demand system and the investor's

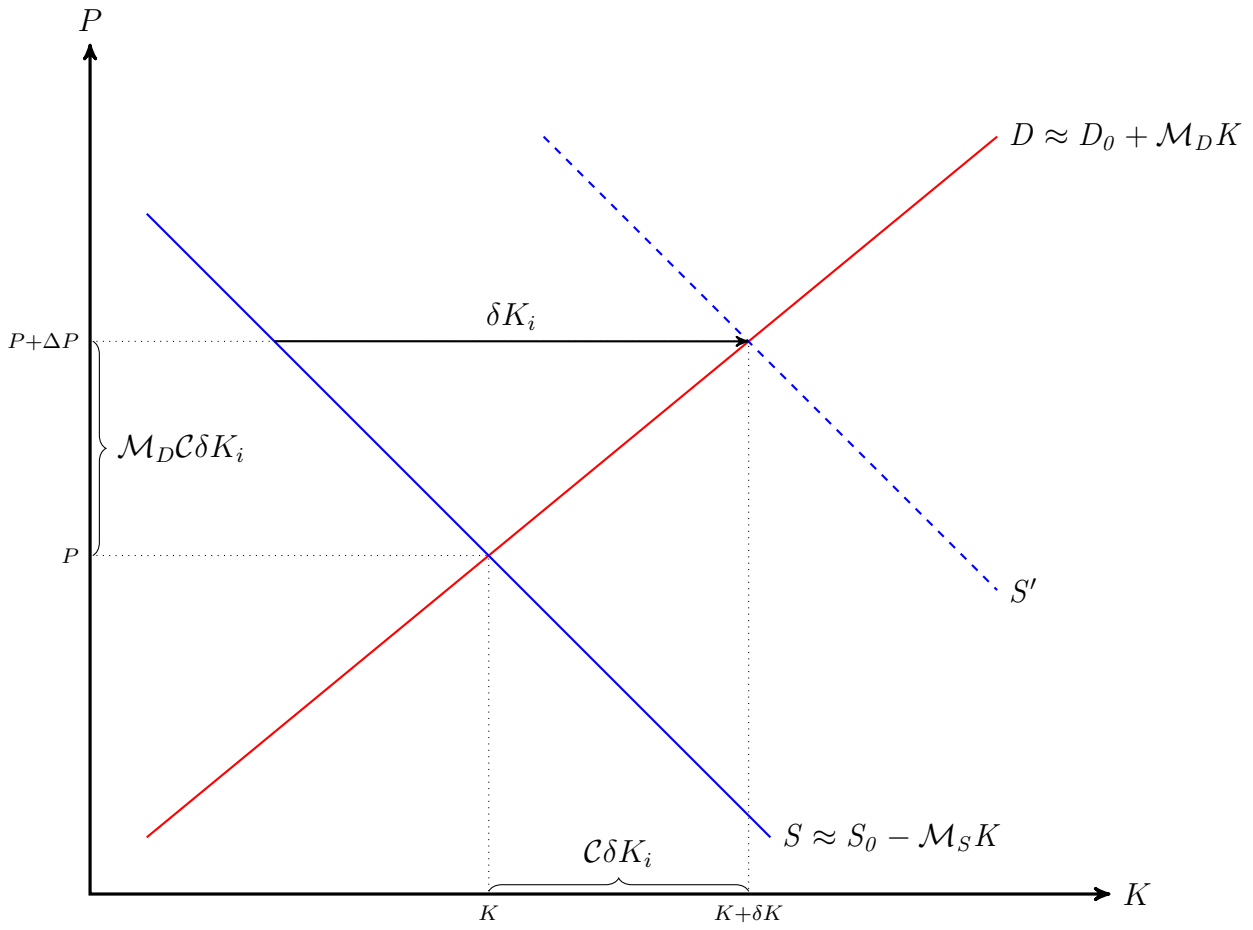


Figure 1. **Impact of a shift in an investor's supply curve**. This stylized diagram illustrates the effect of an individual investor shifting their supply of capital for a single firm. All other investors' supply curves are assumed fixed so that the aggregate investor supply curve moves from S to S' based on the investor's chosen shift δK_i. The point at which the market clears moves along the firm's fixed demand curve D. The equilibrium moves from (K, P) to $(K+\delta K, P+\Delta P)$ with price shift $\Delta P = \mathcal{M}_D \mathcal{C} \delta K_i$ and capital shift $\delta K = \mathcal{C} \delta K_i$, where $\mathcal{C} = \frac{\mathcal{M}_S}{\mathcal{M}_S + \mathcal{M}_D}$ is the contribution multiplier.

marginal contribution to equilibrium outcomes.

To quantify the magnitude and heterogeneity of contribution, I estimate the model building on BCJ's empirical approach and Merton (1987)'s investor supply structure, allowing for cross-sectional heterogeneity in supply and demand parameters. In the calibrated U.S. equity application, the resulting contribution multipliers are economically meaningful and heterogeneous (with substantial variation across firms), consistent with calibrated magnitudes in Favilukis, Garlappi, and Uppal

(2023) and with the modest average real effects estimated in reduced-form studies of investor supply shifts (Berg, Heeb, and Kölbel, 2023; Noh, Oh, and Song, 2023; Briere and Ramelli, 2022).

Using the estimated system, I construct the impact frontier and compare it to an "ESG-efficient" frontier (Pedersen, Fitzgibbons, and Pomorski, 2021) that tilts only on raw scores, as well as intermediate frontiers that partially ignore general-equilibrium spillovers. A key finding is that the impact frontier differs materially from the ESG-efficient frontier; portfolios based only on raw scores can lie far below the frontier and can even imply negative expected impact once real-economy pass-through and cross-firm interactions are taken into account.

This paper contributes to the broader literature on asset demand, demand curves, and production economies (Haddad, Huebner, and Loualiche, 2021; Gabaix and Koijen, 2021; Van der Beck, 2021; Bai and Zhang, 2022; Choi et al., 2022; Gonçalves, Xue, and Zhang, 2020; Bai et al., 2019; Ma, 2019; Belo, Xue, and Zhang, 2013), and to the growing empirical literature on sustainability-related investor preferences and willingness-to-pay (Riedl and Smeets, 2017; Hartzmark and Sussman, 2019; Barber, Morse, and Yasuda, 2021; Bauer, Ruof, and Smeets, 2021; Heeb et al., 2022; Siemroth and Hornuf, 2021; Krueger, Sautner, and Starks, 2020). The model highlights that quantifying investor impact on climate- and nature-related externalities requires jointly modeling investors' supply of capital and firms' demand for capital, especially when the target is real effects over multi-year horizons with more elastic supply and demand.

This paper is organized as follows. Section 2 introduces the supply-and-demand system and its equilibrium when impact preferences are turned off. Section 3 introduces the externality, defines investor impact, and develops the impact frontier when impact preferences are turned on. Section 4 estimates the model, and Section 5 uses the estimates to construct contribution multipliers and impact frontiers. Section 6 discusses limitations and extensions and Section 7 concludes.

2 Baseline Model with Supply and Demand

This section introduces my model for the supply and demand for capital and presents the solution when impact preferences are turned off. The model extends BCJ's supply and demand framework, analyzing firm and investor behavior in terms of their optimal supply and demand schedules rather than equilibrium quantities. This approach is crucial for properly analyzing how changes in investor supply affect firm behavior and the externality when impact preferences are turned on in the next section.

There are I investors indexed by $i \in \{1, \ldots, I\}$, N firms indexed by $n \in \{1, ..., N\}$, and a single period between two times $t \in \{0, 1\}$. Vectors and matrices are in bold and their elements are indicated in parentheses (e.g., $x(n, m)$ is the nth row and mth column of matrix $\boldsymbol{x}$)[2]. The full derivation of the model is presented in the Appendix.

2.1 Firms

At time 0, each firm has installed capital $K_0(n) \geq 0$ (after depreciation) and no other resources or liabilities. The firm's capital is a composite of its tangible and intangible capital. There is one share outstanding for each unit of installed capital and each investor initially holds $K_{i,0}(n)$ shares with $K_0(n) = \sum_{i=1}^{I} K_{i,0}(n)$.

At time 1 each firm generates free cash flow $CF(n) = A(n)K(n)$, where $\boldsymbol{A}$ is a random vector of firm profitabilities with exogenous mean vector $\boldsymbol{a}$ and covariance matrix $\boldsymbol{\Sigma}_{CF} = \boldsymbol{b}_{CF}\boldsymbol{b}'_{CF} + \boldsymbol{\sigma}^2_{idio,CF}$. Here $\boldsymbol{b}_{CF}$ is a vector of sensitivities to a common cash flow factor with unit variance and $\boldsymbol{\sigma}^2_{idio,CF}$ is a vector of idiosyncratic variances. To keep the model self-contained within the single period, I assume the cash flows capture not just profits from production during the period but also the ongoing market values of the installed capital.

[2]Negative row or column numbers indicate the removal of that row or column from the vector or matrix. $\boldsymbol{x}'$ indicates the transpose of $\boldsymbol{x}$. If $\boldsymbol{x}$ is a vector, then $\mathrm{diag}(\boldsymbol{x})$ is the diagonal matrix with $\boldsymbol{x}$ on the diagonal, otherwise it is the vector of diagonal values of the matrix $\boldsymbol{x}$. I denote the identity matrix as $\boldsymbol{I}$ and a vector with the nth element equal to one and other elements equal to zero as $\boldsymbol{e}_n$.

2.2 Investors

Each investor's $t = 0$ wealth is equal to $W_{i,0}$ and the total initial wealth is $W_0 = \sum_{i=1}^{I} W_{i,0}$[3]. They may invest their wealth in $\boldsymbol{K}_i$ shares of the firms in their opportunity set $\mathcal{N}_i \subseteq \{1, ..., N\}$ and in a risk-free asset with exogenous interest rate r_f that is in zero net supply. Their opportunity set may be a subset of the entire investment universe for a variety of underlying reasons, including lack of information, limitations on short sales, divestment commitments, taxes, transaction costs and other frictions. Otherwise, I assume that all investors have the same beliefs, financial preferences and effectively infinitesimal size relative to the market value of each firm. In addition, as in Merton (1987), I assume that all investors have access to an $(N+1)^{st}$ asset that is a forward contract on the cash flow factor and that is in zero net supply.

Each investor has standard mean-variance utility plus a holdings-based preference term,

$$U_i = E[W_{i,1}] - \frac{\gamma}{2W_{i,0}} Var[W_{i,1}] + \boldsymbol{h}_i' \boldsymbol{K}_i, \tag{2}$$

where $W_{i,1}$ is the investor's wealth at time 1, γ defines the level of financial risk aversion (or 'price of risk") and is assumed common across investors, and $\boldsymbol{h}_i$ is a vector of exogenous holdings-based preferences that may be due to non-financial risk, subjective financial beliefs, and ESG preferences[4].

As highlighted by BCJ in an extension of their model, if shifts by active investors cause passive investors to copy or react to those changes (i.e., to continue holding the market portfolio), then the effects of active investors may be leveraged. To account for this, I assume that the set of investors may be partitioned into a subset of active investors, $\mathcal{A}$, and a subset of passive investors, $\mathcal{P}$. Each passive investor seeks to hold $\boldsymbol{K}_{\mathcal{P}} = \boldsymbol{\kappa}_{\mathcal{P}} \boldsymbol{K}_{\mathcal{A}}$, with $\boldsymbol{\kappa}_{\mathcal{P}}$ a diagonal matrix that summarizes the multiple of the active portfolio they choose to hold for each asset. $\boldsymbol{\kappa}_{\mathcal{P}}$ is assumed to be fixed exogenously and may depend, for example, on the relative wealth and risk aversion of the passive investors as in

[3] In the full derivation, I allow each investor's initial wealth to be endogenously determined by their endowed holdings in each firm, given by the vector $\boldsymbol{K_{i,0}}$ plus cash C_i.

[4] See also, the discussion of 'supply shifters' in BCJ.

BCJ, the demand structures explored in Gabaix and Koijen (2021), or the benchmarking intensity of Pavlova and Sikorskaya (2023). In my baseline model I assume $\kappa_{\mathcal{P}}(n) = \frac{W_0}{W_{0,\mathcal{A}}} - 1$ for all assets.

2.3 Equilibrium Supply and Demand

The market clears at the set of asset prices where aggregate demand equals aggregate supply. Similar to Haddad, Huebner, and Loualiche (2021), investors are sophisticated and strategically update their supply curves to be optimal for their own expected utility given their beliefs about aggregate demand and supply[5]. This can be interpreted as firms and investors submitting supply and demand schedules to a neutral market maker, who then clears the market. This is an important but subtle shift from standard models, because in a supply and demand system agents don't have direct control over how much they are able to sell or buy, but only on the schedule they offer to the market.

In general, the equilibrium may not be unique or exist and determining the optimal investor supply may be complex (e.g., for investors that are large or have concentrated initial holdings). However, for the special case of small, uncoordinated investors, the following proposition holds.

Proposition 1. *Let $\boldsymbol{I}_i$ denote the $|\mathcal{N}_i| \times N$ matrix where for each $n \in \mathcal{N}_i$ the corresponding row in $\boldsymbol{I}_i$ is $\boldsymbol{e}_n'$. The optimal supply curve for investor i is*

$$\boldsymbol{P}_i = \frac{1}{1+r_f}\boldsymbol{I}_i(\boldsymbol{a} + \boldsymbol{h}_i) - \boldsymbol{\mathcal{M}}_i\boldsymbol{I}_i\boldsymbol{K}_i, \tag{3}$$

with supply slope matrix

$$\boldsymbol{\mathcal{M}}_i = \frac{1}{1+r_f}\frac{\gamma}{W_{i,0}}\boldsymbol{I}_i\boldsymbol{\Sigma}_{CF}\boldsymbol{I}_i'. \tag{4}$$

Other than the opportunity set restrictions ($\boldsymbol{I}_i$), these are just supply curves for a conventional mean-variance model with a subjective tilt ($\boldsymbol{h}_i$). At the aggregate level, the effect of the heterogeneous

[5]Other papers where investors explicitly optimize their demand schedule include Edmans, Levit, and Schneemeier (2022).

opportunity sets becomes clearer.

Proposition 2. *The aggregate supply curve is*

$$\boldsymbol{P}_S = \boldsymbol{\mu}_S - \boldsymbol{\mathcal{M}}_S \boldsymbol{K}, \tag{5}$$

with $\boldsymbol{\mu}_S = \frac{1}{1+r_f}(\boldsymbol{a} + \tilde{\boldsymbol{h}})$ *and aggregate investor-supply slope matrix*

$$\boldsymbol{\mathcal{M}}_S = \frac{1}{1+r_f}\frac{\gamma}{W_0}\tilde{\boldsymbol{\Sigma}}_{CF}, \tag{6}$$

based on the modified cash flow covariance matrix

$$\tilde{\boldsymbol{\Sigma}}_{CF} = \boldsymbol{b}_{CF}\boldsymbol{b}'_{CF} + \mathrm{diag}(\boldsymbol{q}^{-1})\boldsymbol{\sigma}^2_{idio,CF}, \tag{7}$$

with $q(n) \equiv \sum_{i \in \mathcal{A} \text{ s.t. } n \in \mathcal{N}_i} \frac{W_{i,0}}{W_{0,\mathcal{A}}}$ *the proportion of active investors that have asset* n *in their opportunity sets and modified holdings-based tilt* $\tilde{\boldsymbol{h}} \equiv \frac{1}{1+r_f}\frac{\gamma}{W_{0,\mathcal{A}}}\tilde{\boldsymbol{\Sigma}}_{CF}\sum_{j \in \mathcal{A}} \boldsymbol{I}'_j \boldsymbol{\mathcal{M}}_j^{-1} \boldsymbol{I}_j \boldsymbol{h}_j$.

The proportions $\boldsymbol{q}$ interact with the idiosyncratic cash flow variances to determine aggregate investor supply, in a similar fashion to Merton (1987). Firms with more volatile cash flows (large $b(n)$ or $\sigma^2_{idio,CF}(n)$), or that are in fewer investors' opportunity sets (small $q(n)$), will tend to have higher associated investor-supply slopes. The strong relationship between firm size and demand elasticity found in Haddad, Huebner, and Loualiche (2021) underscores the importance of this variation.

The firm's demand for capital curve is assumed to follow

$$\boldsymbol{P}_D = \boldsymbol{\mu}_D + \boldsymbol{\mathcal{M}}_D(\boldsymbol{K} - \boldsymbol{K}_0), \tag{8}$$

where the slope matrix $\boldsymbol{\mathcal{M}}_D$ is initially assumed to be a diagonal matrix. The parameters of these curves are assumed exogenously determined by firm management on behalf of the incumbent

shareholders to maximize firm market value at $t = 0$. Other stakeholder preferences may also be embedded in these curves, but this is not explored in the present paper.

2.4 Equilibrium Values and Asset Returns

Equilibrium occurs at the market clearing point where aggregate demand equals supply. It is useful to define the 'total slope' matrix $\boldsymbol{\mathcal{M}}_T \equiv \boldsymbol{\mathcal{M}}_S + \boldsymbol{\mathcal{M}}_D$.

The market clears and equilibrium is achieved at

$$\boldsymbol{K} = \boldsymbol{\mathcal{M}}_T^{-1}(\boldsymbol{\mu}_S - \boldsymbol{\mu}_D + \boldsymbol{\mathcal{M}}_D \boldsymbol{K}_0), \tag{9}$$

$$\boldsymbol{P} = \boldsymbol{\mu}_S - \boldsymbol{\mathcal{M}}_S \boldsymbol{K}, \tag{10}$$

As expected asset returns are given by $E[R(n)] - r_f = \frac{a(n)}{P(n)} - (1 + r_f)$, they depend on the supply and demand slope matrices via $\boldsymbol{P}$. The following proposition shows how these expected returns differ substantially from the CAPM (Sharpe, 1964).

Proposition 3. *The expected asset returns may be written as*

$$E[R(n)] - r_f = \beta_M(n)(E[R_M] - r_f) + (\lambda(n) - \beta_M(n)\lambda_M) \tag{11}$$

where $\beta_M(n)$ is the standard beta of the asset to the market return R_M, and

$$\lambda(n) \equiv P(n)^{-1}\Big(\frac{\gamma}{W_0}(q(n)^{-1} - 1)\sigma^2_{idio,CF}(n)K(n) - \tilde{h}(n)\Big), \tag{12}$$

is a non-market risk premium that is in addition to the premium for systematic risk and λ_M represents the total non-market risk premium associated with the market portfolio.

This proposition generalizes the results of Merton (1987)[6]. There are two components to the non-

[6]His results may be recovered by setting $\boldsymbol{\Sigma}_{ICS} = W_{0,\mathcal{P}} = \boldsymbol{\kappa}_{\mathcal{P}} = \boldsymbol{h}_i = 0$. In this case, $\boldsymbol{P} = \mathbf{1}$, so $\boldsymbol{\lambda} = \gamma(\mathrm{diag}(\boldsymbol{q}^{-1}) - \boldsymbol{I})\boldsymbol{\sigma}^2_{idio,CF}\boldsymbol{x}$, with $\boldsymbol{x} = \boldsymbol{K}/W_0 = \boldsymbol{V}/W_0$ matching his equations (23), (19) and (15).

market risk premium. First, as in Merton (1987), there is the term reflecting the difference between $\tilde{\boldsymbol{\Sigma}}_{CF}$ and $\boldsymbol{\Sigma}_{CF}$ due to investor's heterogeneous opportunity sets. This aligns with the findings of Dello Preite et al. (2022), who demonstrate that asset pricing models that allow compensation for non-market risk outperform traditional factor models. Second, assets that are subjectively avoided for holdings-based reasons (negative $\tilde{h}$) will, all else equal, have higher CAPM alphas. Equation (9) means such firms will be smaller (and less correlated with the market). BCJ highlight that the betting-against-beta and betting-against-correlation anomalies can thus be understood as originating from cross-sectional heterogeneity in these characteristics.

Note that equation (12) is not in closed form as it depends on equations (9) and (10). However, this only emphasizes that as long as $q(n) \neq 1$ and $\tilde{h}(n) \neq 0$ then $\lambda(n)$ will be non-zero and vary in complex ways over the cross-section. In particular, at first glance equation (12) appears linear in $K(n)$, so that for a fixed $P(n)$ larger firms will have higher expected returns. Suppose, however, in line with the estimates in Merton (1987), that large, widely-held firms have $q(n) \approx 1$ while smaller, lesser-known firms have $q(n) \approx 0.03$. Then large firms will have $\lambda(n) \approx 0$, while smaller firms will have $q_{small}^{-1}(n) - 1 \approx 30$ resulting in $\lambda > 0$.

3 Equilibrium with Investor Impact

In this section, I add an externality to the model, formally define 'impact', and turn on impact preferences.

3.1 Externality

The aggregate externality is a function $G(\boldsymbol{K})$ of the vector of installed capitals $\boldsymbol{K}$. In the region of the equilibrium, I assume that $g(n) \equiv \frac{\partial G}{\partial K(n)}$ is an exogenous coefficient that defines the 'impact productivity' of firm n, so $G(\boldsymbol{K}) \equiv \sum_{n=1}^{N} g(n)K(n)$ is approximately a simple linear function. In climate and nature finance applications, G may represent an outcome such as cumulative emissions,

aggregate biodiversity impact, or a composite externality index evaluated over a multi-year horizon.

The impact productivities, $g(n)$, are exogenous and defined so that G is positive in the direction that is considered 'good' (so, for example, positive CO2 emissions produce negative g, lowering G). G could be a social or environmental externality, the aggregate consumer surplus (Allcott et al., 2022), or even purely financial values like the expected long-term market return—the defining feature of G is that it is an aggregate quantity that depends on $\boldsymbol{K}$. The timing of the externality does not matter in this single-period model; it can be interpreted as summarizing real effects that materialize over multi-year horizons.

3.2 Investor Impact

Before defining impact preferences, it is useful to define 'investor impact' and the contribution multiplier matrix, $\boldsymbol{\mathcal{C}}$.

Definition 1. *Investor Impact:* The change in the equilibrium value of the externality, G, due to a marginal shift $\delta \boldsymbol{K}_i$ in an investor's supply schedule.

Proposition 4. *The investor impact due to supply shift $\delta \boldsymbol{K}_i$ is, to first order, given by*

$$II_i \equiv \boldsymbol{g}_i' \boldsymbol{\mathcal{C}} \delta \boldsymbol{K_i}. \tag{13}$$

with 'contribution multiplier' matrix $\boldsymbol{\mathcal{C}} \equiv (\boldsymbol{M}_D + \boldsymbol{M}_S)^{-1} \boldsymbol{M}_S \left(\boldsymbol{I}_N - \frac{\boldsymbol{BD}}{\boldsymbol{DB}}\right)(\boldsymbol{I} + \boldsymbol{\kappa}_{\mathcal{P}})$ with $\boldsymbol{D} \equiv ((\boldsymbol{K} - \boldsymbol{K}_0)' \boldsymbol{M}_D + \boldsymbol{P}')(\boldsymbol{M}_D + \boldsymbol{M}_S)^{-1} \boldsymbol{M}_S$ and $\boldsymbol{B} \equiv \frac{1}{1+r_f} \boldsymbol{M}_S^{-1} \boldsymbol{P}$. As I show in the appendix, $\boldsymbol{\mathcal{C}}(n, m)$ represents the elasticity of $K(n)$ with respect to an increase in $K_i(m)$ (i.e., the contribution of a shift in supply for m to the equilibrium amount of n).

In the sequence of matrix multiplications that determines the contribution multiplier matrix, $\boldsymbol{M}_S$ can be thought of as converting investor supply shifts into price shocks, while $(\boldsymbol{M}_D + \boldsymbol{M}_S)^{-1}$ converts the price shocks back into changes in capital. $\left(\boldsymbol{I}_N - \frac{\boldsymbol{BD}}{\boldsymbol{DB}}\right)$ is a general equilibrium term to enforce that wealth is kept within the market, so prices adjust to ensure the risk-free asset remains in

zero net supply—asset allocation effects, shifting capital between markets, are a separate channel and beyond the scope of this paper. $(\boldsymbol{I} + \boldsymbol{\kappa}_{\mathcal{P}})$ represents the effect of passive investment.

Note that the multipliers depend only on the relative values of the slope matrices, not their levels. They are zero when either firm demand is fully inelastic ($\mathcal{M}_D(n,n) \to \infty$) or investor supply is highly elastic ($\mathcal{M}_S(n,n) \to 0$). These are explicit or implicit assumptions in many models. However, when neither of these assumptions are imposed investor supply shifts have the potential to generate significant investor impact[7].

In my baseline model $\boldsymbol{\beta}_{T,n} = 0$ and, assuming $\Sigma_{ICS}(n,n)$ is large relative to $\Sigma_{CF}(n,n)$, the additionality multiplier is

$$\mathcal{C}^A(n) \equiv \mathcal{C}(n,n) \approx \frac{W_0}{W_{0,\mathcal{A}}} \frac{b_{CF}^2(n) + \frac{\sigma_{idio,CF}^2(n)}{q(n)}}{b_{CF}^2(n) + \frac{\sigma_{idio,CF}^2(n)}{q(n)} + \mathcal{M}_D(n,n)}. \tag{14}$$

This shows that, all else equal, more neglected firms with smaller values of $q(n)$ will have higher contribution multipliers.

3.3 Impact preferences

Now suppose investors have the mean-variance preferences U_i from equation (2), plus a taste for the externality,

$$\tilde{U}_i = U_i + \gamma_i^g G. \tag{15}$$

The 'price of impact' (γ_i^g) defines the investor's marginal rate of substitution between the externality and the financial good at $t = 0$ and is assumed exogenous.

[7]That said, contribution multipliers are unrelated to ideas about 'attribution' where one might seek to 'fairly' allocate one hundred percent of responsibility for the externality across the investors involved—contribution multipliers are about marginal impact and will typically be significantly less than one.

Proposition 5. *Turning on impact preferences has the following effects. It shifts the individual investor supply curve from $\boldsymbol{P}_i$ in equation (3) to*

$$\tilde{\boldsymbol{P}}_i = \boldsymbol{P}_i + \boldsymbol{I}_i \boldsymbol{\mathcal{C}}' \gamma_i^g \boldsymbol{g}, \tag{16}$$

generating an impact-tilt portfolio of $\delta \boldsymbol{K}_i = \gamma_i^g \boldsymbol{I}_i \boldsymbol{\mathcal{M}}_i^{-1} \boldsymbol{I}_i \boldsymbol{\mathcal{C}}' \boldsymbol{g}$ and investor impact of $II_i = \gamma_i^g \boldsymbol{g}_i' \boldsymbol{\mathcal{C}} \boldsymbol{I}_i \boldsymbol{\mathcal{M}}_i^{-1} \boldsymbol{I}_i \boldsymbol{\mathcal{C}}' \boldsymbol{g}$.

In short, comparison of equations (13) and (16), shows that each investor acts as if the assets produce impact 'value' and offers 'impact returns' of

$$\boldsymbol{R}_{\boldsymbol{i}}^{(\boldsymbol{I})} \equiv \gamma_i^g \operatorname{diag}(\boldsymbol{P}^{-1}) \boldsymbol{\mathcal{C}} \boldsymbol{g}, \tag{17}$$

in addition to financial returns and any holdings-based tilts. While investors are assumed to be effectively infinitesimal and uncoordinated, so their individual investor impact will be small, they still tilt their portfolios based on these impact returns because there is zero marginal cost to do so (starting from the baseline equilibrium).

3.4 Impact Frontier

The concept of the 'impact frontier' is a key tool for understanding the trade-off between certainty-equivalent returns and investor impact.

Definition 2. *Impact Frontier:* Investor i's 'impact frontier' is the Pareto frontier between certainty-equivalent returns and the expected effect on an externality (investor impact).

Visually, as in Figure 2, an investor's impact frontier may be represented as the curve of their certainty-equivalent return (financial utility) versus their investor impact II_i as their price of impact varies.

Proposition 6. *The impact frontier is, considering only the parts of U_i that vary substantially with*

$\boldsymbol{K}_i$,

$$U_i^{(IF)} \approx U_{i,0} - \frac{1+r_f}{2} \frac{1}{(\boldsymbol{g}_i' \boldsymbol{\mathcal{C}} \boldsymbol{\mathcal{M}}_i^{-1} \boldsymbol{\mathcal{C}}' \boldsymbol{g}_i)} II_i^2, \tag{18}$$

with $U_{i,0}$ the value of U_i for $\gamma_i^g = 0$ and II_i as in Proposition 5.

While this structure is the same for all investors, the specific parameters that determine these curves may vary across investors (e.g., the desired level of II_i, the value of $\boldsymbol{\mathcal{M}}_i$). As in Merton (1987), the market portfolio will not be mean-variance efficient for investors that have access to more securities than the average market participant, and hence for these investors it will also not be on the impact frontier.

At each point on the impact frontier, the absolute value of the slope of the tangent line (the 'price of impact line') is equal to the price of impact, γ_i^g, associated with the portfolio at that point.

In practice, investors will have (at best) noisy estimates of the contribution multiplier matrix and impact productivities. Thus, it is important to consider the effect of errors in these assessments relative to the ideal with perfect information. The following proposition tackles this question by wrapping all potential errors into a flawed contribution multiplier matrix $\boldsymbol{\mathcal{C}}^*$.

Proposition 7. *The 'inefficient' impact frontier traced out by the portfolios of an investor that mistakenly shifts their supply based on contribution multiplier matrix $\boldsymbol{\mathcal{C}}^* \neq \boldsymbol{\mathcal{C}}$ is*

$$U_i^* \approx U_{i,0} - \frac{1+r_f}{2} \frac{\boldsymbol{g}_i' \boldsymbol{\mathcal{C}}^* \boldsymbol{\mathcal{M}}_i^{-1} \boldsymbol{\mathcal{C}}^{*\prime} \boldsymbol{g}_i}{(\boldsymbol{g}_i' \boldsymbol{\mathcal{C}} \boldsymbol{\mathcal{M}}_i^{-1} \boldsymbol{\mathcal{C}}^{*\prime} \boldsymbol{g}_i)^2} II_i^2 \leq U_i^{(IF)}. \tag{19}$$

The net utility gain from an 'inefficient' tilt, as in the above proposition, will be proportional to $\boldsymbol{g}'(\boldsymbol{\mathcal{C}} - \frac{1}{2}\boldsymbol{\mathcal{C}}^*)\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{C}}^{*\prime}\boldsymbol{g}$. Thus, tilts that are based on contribution estimates that are too optimistic, $\boldsymbol{\mathcal{C}}^* > 2\boldsymbol{\mathcal{C}}$ will result in a net decrease in the investor's utility. This means it could be rational for consequentialist investors who are uncertain about their contribution to be conservative in how they tilt their portfolios.

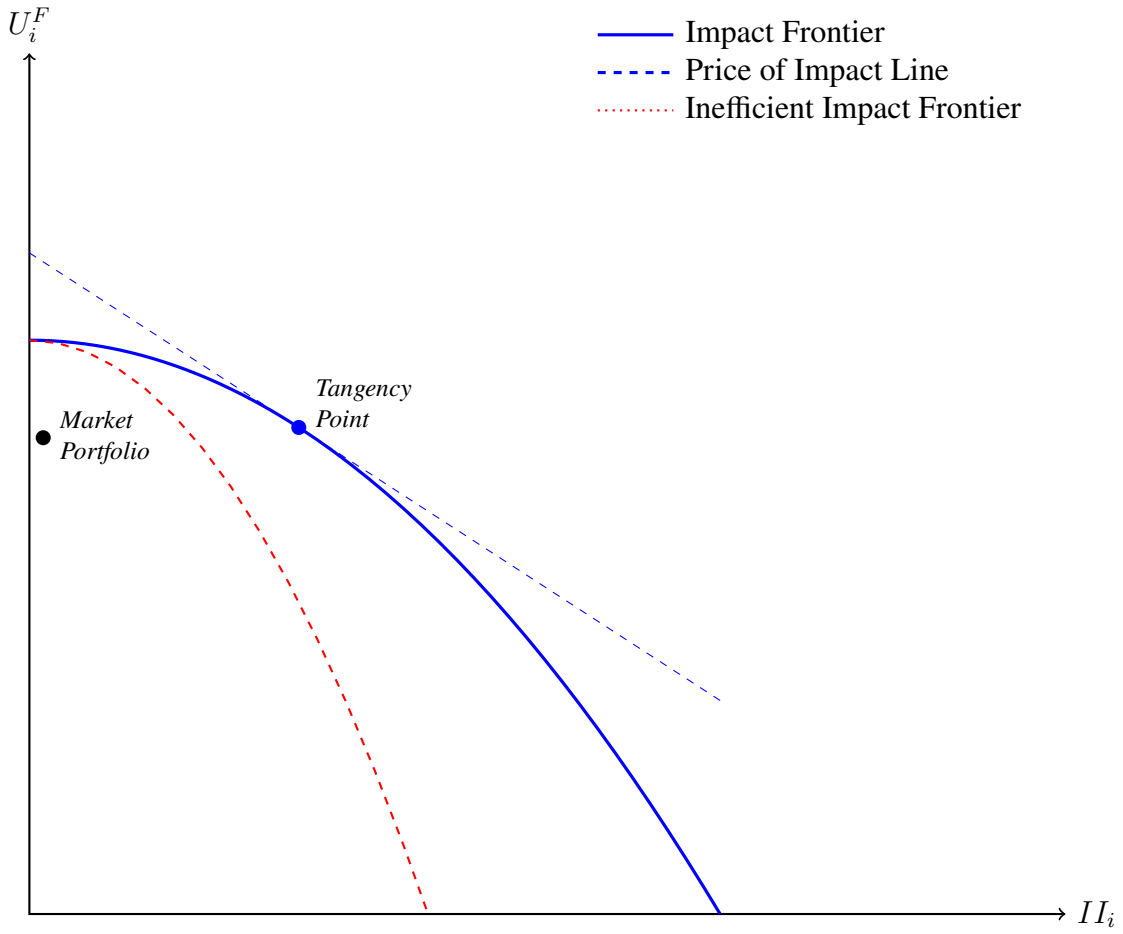


Figure 2. **The Impact Frontier**. An investor's 'impact frontier' is the set of optimal portfolios, expressed in terms of the associated investor impact (II_i) and certainty-equivalent returns (financial utility, U_i^F), as the investor's price of impact is increased from zero. At each point on the impact frontier, the associated tangent 'Price of Impact Line' has slope equal to the price of impact. In general, the market portfolio need not be on the frontier. The 'inefficient' impact frontier derived from an incorrect contribution multiplier matrix is strictly below the impact frontier.

4 Empirical Estimation

This section details the empirical estimation of the supply and demand system. The empirical strategy builds on the joint estimation of supply and demand curves pioneered by BCJ while incorporating Merton (1987)'s insights on heterogeneous investor opportunity sets. Section 4.1 describes the data, sample construction, and key variables. Section 4.2 presents the econometric

specification and discusses identification. Section 4.3 reports the estimation results from Ordinary Least Squares (OLS), Two-Stage Least Squares (2SLS), and the preferred Three-Stage Least Squares (3SLS) specifications.

4.1 Data and Variable Construction

My analysis utilizes data for U.S. stocks from the merged CSP-Compustat database covering the period 1999-2024. Following BCJ, I construct an unbalanced panel focusing on cross-sections at year-end 2004, 2009, 2014, and 2019. Standard filters are applied: I exclude financial firms (SIC 6000-6999) and utility firms (SIC 4900-4949) and require at least one year of return data. Firm-level externality-score proxies are sourced from MSCI and are used throughout as *raw externality scores* for illustration (e.g., MSCI pillar scores). The construction of key variables largely follows BCJ, unless otherwise noted.

My primary dependent variable is the forward-looking Sharpe ratio ($\lambda_{n,t}$), estimated using monthly dollar excess returns over the five years following year t. Other variables related to risk quantities and firm decisions are calculated over the same window and using dollar returns, including the external component of market correlation ($\rho^{ext}_{M,n,t}$) and the internal component of market correlation ($\rho^{int}_{M,n,t}$).

In line with the investment CAPM literature (Zhang, 2017), I use log book equity growth as the primary measure of firms' investment decisions and to naturally manage outliers. This practical choice aligns with q-theory's emphasis on actual capital deployment rather than market valuations. The log 5-year book equity growth ($\ln(BEG_{n,t})$) is calculated over the period $t-5$ to t and winsorized at the 1st and 99th percentiles to further mitigate the influence of extreme outliers. Firm-specific characteristics such as profitability, cash flow volatility, and tangible intensity are derived from Compustat data for fiscal years ending in t or based on multi-year windows ending then. Firms are assigned to one of three size groups ($G_k, k = 1, 2, 3$, with higher k indicating greater size) based on market capitalization breakpoints within the cross-section of each year t,

with the breakpoints chosen to align with the Merton groups in Table V. This grouping focuses differentiation on the larger firms in the sample.

Table I presents summary statistics for the key variables used in the estimation for the final sample comprising 3,046 stock-year observations after applying all filters and variable requirements.

4.2 Econometric Specification and Identification

I consider a system of two simultaneous equations, representing investor supply and firm demand. The equations are based on the theoretical price-quantity relations given by equations (5) and (8), but as in BCJ translated into equations with the Sharpe ratio ($\lambda_{n,t}$) as the representative of price. The investor supply equation also relates to equations (11) and (12).

The investor supply equation models the required Sharpe ratio as a function of priced risk quantities and investor-specific supply shifters:

$$\lambda_{n,t} = \boldsymbol{x}'_{S,n,t}\boldsymbol{\beta}_S + \Delta_{S,sys}\rho^{ext}_{M,n,t} + \sum_{k=1}^{3} \Delta_{S,nonmkt,k}\mathbb{I}(n \in G_k)\rho^{int}_{M,n,t} + u_{S,n,t} \tag{20}$$

Here, $\rho^{ext}_{M,n,t}$ represents the estimated quantity of systematic risk (a firm's correlation with the market, excluding correlation with its idiosyncratic risk), while the interaction term $\mathbb{I}(n \in G_k)\rho^{int}_{M,n,t}$ captures the quantity of non-market risk, with its pricing potentially varying across three size groups (G_k) as motivated by Merton (1987). Both types of risk quantities are treated as endogenous. The vector $\boldsymbol{x}_{S,n,t}$ contains exogenous investor supply shifters (social score, governance score) and controls (Industry and Year fixed effects). $\Delta_{S,nonmkt,k}$ are the key investor-supply slope parameters that determine the contribution multipliers.

The firm demand equation models the Sharpe ratio in relation to the firm's capital demand decision, proxied by its growth, and firm-specific demand shifters:

$$\lambda_{n,t} = \boldsymbol{x}'_{D,n,t}\boldsymbol{\beta}_D + \Delta_D \ln(BEG_{n,t}) + u_{D,n,t} \tag{21}$$

	Mean	SD
Market portfolio		
Risk premium (% per year) (1)	9.116	
Volatility (% per year) (1)	15.157	
Sharpe ratio	0.601	
Firms		
Risk premium (% per year) (2)	11.779	17.773
Volatility (% per year) (2)	37.323	17.565
Sharpe ratio (forward 5Y)	0.384	0.440
Market correlation	0.510	0.169
Market cap ($ billion)	13.999	49.083
Book equity ($ billion)	4.668	16.859
log Book equity growth (5Y)	0.455	0.726
Environment score	−1.962	1.408
Social score	−2.594	1.265
Governance score	−0.905	0.832
Business equity correlation	0.010	0.020
Profitability	2.718	3.042
Cash flow volatility	4.502	4.677
Tangible intensity	14.790	13.124

Table I. **Summary Statistics for the Full Panel.** This table reports summary statistics on the panel of U.S. stocks. I construct the panel from the merged CRSP-Compustat database, to which I apply the filters described in the main text. The panel consists of four cross-sections corresponding to December 31 of years 2004, 2009, 2014, and 2019 and contains 1,396 stocks and 3,046 stock-year observations. I compute firm statistics on December 31 of year t as follows. Market equity and book equity are assessed on the day. Risk capital is the stock's annualized volatility of monthly dollar excess returns between the beginning of year t-4 and the end of year t. Market correlation is also estimated from monthly excess returns from t-4 to t. Risk premia, volatility, and Sharpe ratio are estimated using monthly stock returns between t and t+5. Average profitability per unit of risk (profitability) is calculated from profitability data between t- 9 and t. Return on assets (ROA) is measured as the ratio of the firm's quarterly income to its lagged book asset value. Cash flow volatility is the volatility of ROA, and profitability per unit of risk is the ratio of average ROA to cash flow volatility. Tangible intensity is the ratio of tangible to total capital per unit of risk at t, interacted with a dummy variable equal to unity if the firm has non-zero debt at t. Business equity correlation is the correlation between the stock's quarterly return and the growth of aggregate business equity held by households from t-4 to t. Reported statistics are equal-weighted averages of firm-year variables.

The log 5-year book equity growth, $\ln(BEG_{n,t})$, serves as the endogenous proxy for the quantity of assets supplied or the firm's investment decision. The vector $\boldsymbol{x}_{D,n,t}$ includes exogenous firm demand shifters (profitability, tangible intensity, lagged Sharpe ratio) and controls (Industry and Year fixed effects). The lagged Sharpe ratio is included to reflect that it is changes in the cost of capital offered to the firm that drive changes in firm financial size, not levels. Δ_D is the key demand slope parameter.

Identification is based on instrumental variables. Following BCJ, exogenous shifters from one side serve as excluded instruments for the endogenous variable(s) on the other side. For the investor supply equation (20), the endogenous risk quantities ($\rho^{ext}_{M,n,t}$ and the $\rho^{int}_{M,n,t}$ interactions) are instrumented using the firm demand shifters (profitability, cash flow volatility, tangible intensity) and the lagged Sharpe ratio. For the firm demand equation (21), the endogenous growth variable ($\ln(BEG_{n,t})$) is instrumented using the investor supply shifters (ESG pillar scores, correlation with private business equity) and size group indicators. Prior analysis using 2SLS confirmed the strength and validity (based on Sargan tests) of these instrument sets for the chosen specifications, supporting the use of IV methods.

4.3 Estimation Results

The system is estimated using OLS, 2SLS, and 3SLS. OLS estimates, while potentially biased due to endogeneity, provide a baseline. 2SLS provides consistent estimates for each equation separately. 3SLS estimates the system jointly, leveraging the correlation between the equation residuals for potentially greater efficiency.

Table II presents the OLS and 2SLS estimates for the core specifications. The OLS results (Columns 1 and 2) show significant correlations, but the estimated slopes may be biased. For instance, the OLS slope for log book equity growth in the supply equation is negative and significant, while the OLS slopes for the risk quantities in the demand equation are significant but potentially conflate supply and demand effects. The 2SLS results (Columns 3 and 4) apply the IV strategy.

Diagnostics for these specific 2SLS regressions confirm instrument relevance and validity, and Wu-Hausman tests indicate the presence of endogeneity bias in OLS. The 2SLS estimates show plausible signs for the slopes, although statistical significance is borderline for some key coefficients, notably the supply slope.

Table III presents the results from my preferred 3SLS estimation of the system defined by equations (20) and (21). The 3SLS estimates benefit from the efficiency gains from the high cross-equation residual correlation (0.91).

In the firm demand equation, the coefficient on log book equity growth (Δ_D) is negative and statistically significant. This negative slope suggests firms scale back growth plans when faced with a higher cost of capital (required Sharpe ratio). The lagged Sharpe ratio remains a strong positive predictor, indicating persistence.

In the investor supply-of-capital equation, the coefficient on external market correlation ($\Delta_{S,sys}$) is positive and highly significant. This confirms that investors require higher compensation for systematic risk. The coefficients for the non-market risk interactions ($\Delta_{S,nonmkt,k}$) decrease exponentially from the smallest size group to the largest and all are statistically significant. Their positive signs and decreasing magnitude with size strongly support the Merton (1987) hypothesis that non-market risk is priced due to market segmentation, with the effect diminishing for larger, more widely held firms. Results with alternative size groups, shown in Table IV, are qualitatively similar (albeit with weaker statistical significance).

Table V compares my results to Merton's original analysis where he sorted his 1985 sample of 1,387 U.S. firms into 10 decile groups and used the variation in market value per registered shareholder to form rough estimates of the relative size of the investor base of each group. That is, the ratio of the percentage of investors with a firm in their opportunity sets to the same percentage for firms in the largest group. A similar analysis can be done by comparing the ratios between my estimates of $\Delta_{S,nonmkt,k}$. The results are remarkably consistent with Merton's prescient analysis and consistent with the hypothesis that smaller, less widely held stocks should command a premium for non-market risk.

	OLS		2SLS	
	Supply	Demand	Supply	Demand
Investor Supply				
Market correlation (external)	0.488^{***}		1.110^{***}	
	(0.044)		(0.265)	
Social score	0.004		-0.012	
	(0.006)		(0.009)	
Idio. risk $\times$ Size group 1	311.237^{***}		26.386	
	(21.356)		(113.381)	
Idio. risk $\times$ Size group 2	93.254^{***}		4.984	
	(8.934)		(26.991)	
Idio. risk $\times$ Size group 3	16.359^{***}		7.131	
	(2.028)		(5.712)	
Firm Demand				
Log(Book equity growth 5Y)		-0.064^{***}		-0.263^{***}
		(0.010)		(0.071)
Sharpe ratio (lagged)		0.064^{***}		0.169^{***}
		(0.019)		(0.043)
Number of observations	3046	3046	3046	3046
F statistic	70.356^{***}	56.410^{***}	43.470^{***}	47.515^{***}
2SLS Diagnostics				
Exogenous instruments (Sargan)			0.647	24.163
Exogenous OLS errors (Wu-Hausman)			8.347^{***}	9.099

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table II. **OLS and 2SLS Estimates of the Supply and Demand System.** This table reports estimates of supply and demand schedules obtained on an unbalanced panel of 1396 U.S. firms observed on December 31 of years 2004, 2009, 2014, and 2019. Columns 1 and 2 report OLS estimates, columns 3 and 4 report 2SLS estimates. All regressions include time and industry fixed effects.

	(1)	(2)	(3)	(4)
Investor Supply				
Market correlation (external)	0.619^{***}	0.569^{***}	1.731^{***}	1.611^{**}
	(0.195)	(0.181)	(0.643)	(0.645)
Idio. risk $\times$ Size group 1	219.095^{**}	221.297^{**}	577.836	513.250
	(97.981)	(98.729)	(425.920)	(428.235)
Idio. risk $\times$ Size group 2	45.370^{**}	44.671^{**}	109.714	84.266
	(21.642)	(21.374)	(83.978)	(84.056)
Idio. risk $\times$ Size group 3	10.900^{**}	10.618^{**}	26.346^{*}	22.337
	(4.369)	(4.343)	(15.844)	(15.888)
Social score	-0.009			
	(0.006)			
Firm Demand				
Log(Book equity growth 5Y)	-0.142^{***}	-0.129^{***}	-0.596^{***}	-0.549^{**}
	(0.049)	(0.049)	(0.226)	(0.220)
Sharpe ratio (lagged)	0.119^{***}	0.110^{***}	0.501^{***}	
	(0.036)	(0.036)	(0.176)	
Instruments				
Social Score	Yes	No	No	No
Profitability	Yes	Yes	No	No
Cash flow volatility	Yes	Yes	Yes	Yes
Tangible Intensity	Yes	Yes	No	No
Lagged Sharpe Ratio	Yes	Yes	No	No
Number of observations	3046	3046	3046	3046
F statistic	247.343^{***}	256.973^{***}	233.345^{***}	247.223^{***}

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table III. **3SLS Estimates of the Supply and Demand System.** This table reports estimates of supply and demand schedules obtained by three-stage least squares on an unbalanced panel of U.S. firms observed on December 31 of years 2004, 2009, 2014, and 2019. Columns 1 is the baseline model. Column 2 drops the social score. Column 3 also drops profitability, cash flow volatility, tangible intensity and lagged Sharpe ratio as instruments. Column 4 drops lagged Sharpe ratio from the demand equation. All regressions include time and industry fixed effects.

	(1)	(2)	(3)
Investor Supply			
Market correlation (external)	0.673^{***}	0.931^{**}	0.838^{*}
	(0.221)	(0.370)	(0.435)
Idio. risk $\times$ Size bucket 1	107.455	492.590^{*}	910.861
	(110.295)	(259.300)	(556.835)
Idio. risk $\times$ Size bucket 2	31.492	104.847	368.183
	(24.707)	(76.581)	(289.823)
Idio. risk $\times$ Size bucket 3	13.517^{***}	64.693^{*}	210.012
	(5.153)	(39.285)	(176.222)
Idio. risk $\times$ Size bucket 4		33.300	83.288
		(21.640)	(104.927)
Idio. risk $\times$ Size bucket 5		10.899^{**}	57.582
		(5.350)	(64.040)
Idio. risk $\times$ Size bucket 6			45.507
			(39.455)
Idio. risk $\times$ Size bucket 7			23.000
			(21.778)
Idio. risk $\times$ Size bucket 8			7.893
			(5.284)
Social score	-0.002	-0.013	-0.012
	(0.007)	(0.009)	(0.010)
Firm Demand			
Log(Book equity growth 5Y)	-0.238^{***}	-0.130^{**}	-0.105^{**}
	(0.069)	(0.050)	(0.047)
Industry book equity growth 5Y (external)	-1.008^{***}		
	(0.319)		
Sharpe ratio (lagged)	0.212^{***}	0.103^{***}	0.082^{***}
	(0.051)	(0.034)	(0.030)
Instruments			
Social Score	Yes	Yes	Yes
Profitability	Yes	Yes	Yes
Cash flow volatility	Yes	Yes	Yes
Tangible Intensity	Yes	Yes	Yes
Lagged Sharpe Ratio	Yes	Yes	Yes
Number of observations	3046	3046	3046
F statistic	231.011^{***}	232.770^{***}	213.523^{***}

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table IV. **3SLS Estimates of the Supply and Demand System.** This table reports estimates of supply and demand schedules obtained by three-stage least squares on an unbalanced panel of U.S. firms observed on December 31 of years 2004, 2009, 2014, and 2019. Columns 1 adds the five year growth rate of book equity for other firms in the same industry. Columns 2 and 3 use alternative, finer-grained size buckets (presented alongside, but different from, the baseline buckets to conserve space). All regressions include time and industry fixed effects.

Merton (1987)			Baseline Model		Extended Model	
Size decile	% Total Market Value	Inferred Relative Size of Investor Base	% Total Market Value	Inferred Relative Size of Investor Base	% Total Market Value	Inferred Relative Size of Investor Base
1	0.1	0.02	12.7	0.05	0.4	0.009
2	0.3	0.03				
3	0.7	0.04			1.8	0.02
4	1.2	0.05				
5	2.0	0.06			2.0	0.04
6	3.2	0.09			3.2	0.09
7	5.3	0.13			5.3	0.14
8	8.7	0.18	23.8	0.24	8.7	0.17
9	15.1	0.34			15.1	0.34
10	63.5	1	63.5	1	63.5	1

Table V. **Relative size of investor base**. The ten rows of decile groups, and the '% Total Market Value' and 'Relative Size of Investor Base' columns are from Table I in Merton (1987). The relative size of the investor base is an estimate of the proportion of investors that have a stock in their opportunity sets. The size groups in the estimated model are calibrated to represent the same portion of total market value as the respective combination of the Merton groups. The relative size of the investor based implied by the model is inferred from the size interaction terms in Table III.

These 3SLS results provide a set of statistically significant and theoretically coherent estimates for the key supply and demand slopes. They form the basis for calculating the estimated contribution multiplier matrix and analyzing the impact frontier in the subsequent sections. While the statistical precision of some estimates is modest, the qualitative pattern of results is robust across specifications and consistent with theoretical predictions. The primary contribution of this paper is identifying the role of contribution multipliers in determining investor impact, rather than precisely estimating specific asset pricing factors. The empirical results should be viewed as a first step toward quantifying these relationships, with future research refining these estimates using more data and more elaborate empirical specifications.

4.4 Observed Supply Shifts for Raw Externality-Score Proxies

Before constructing the impact frontier, it is useful to benchmark the magnitude of investor supply shifts associated with commonly used sustainability-related scores. This provides context for interpreting the impact frontier results in the following section, as it indicates where on the frontier investors might seek to operate in practice.

Empirical estimates of the supply shifts associated with popular sustainability-related metrics are summarized in Table VI and discussed below. To enable comparison between results from different studies, the estimates are presented as implied 'impact returns' for the average firm in my model. Furthermore, given a supply coefficient for the overall market, it must be the case that some investors have even larger supply coefficients. To illustrate this in a way that is consistent across the metrics examined in this section, I also report results for supply coefficients that are five times larger. Such larger coefficients are still well within the ranges reported by Koijen, Richmond, and Yogo (2024) (KRY) and Noh, Oh, and Song (2023) (NOS).

Social score. As part of their study, BCJ estimate that a one-standard deviation improvement in a firm's MSCI 'Social' (S) score reduces its equilibrium Sharpe ratio by 0.025. For the average firm in their cross-section, with a return volatility of 36%, such a decrease corresponds to a 0.9 percentage point absolute decrease in investor's required returns. Their findings indicate that changes in nonpecuniary characteristics can indeed have an economically significant impact on firms.

Environmental score. Using an approach focused on leveraging investor holdings data, but using a reasonably similar U.S. data set to the above estimates, KRY estimate the shift in investor supply curves from a one standard deviation increase in firms' Sustainalytics environmental (E) score. They report estimated supply coefficients for each institution in their sample, with the estimates ranging from a roughly -40% decline in supply to a 25% increase in supply (i.e., percentage increase in the target weight in the institution's portfolio). As a representative value I use the average coefficient of roughly 5% for households (the largest class of investor in the sample)[8]. Based on the Sharpe

[8]This is a rough estimate as this result is only reported in a histogram chart, not a table. Passive investors, long-term

ratio of the average firm in the BCJ panel, a 5% increase in supply corresponds to a 0.024 change in Sharpe ratio and a 0.9 percentage point shift in return. These values are remarkably similar to those for the social score, suggesting a consistent pattern in how environmental and social characteristics affect investor supply.

Emissions intensity and adjusted environmental score. NOS follow an approach similar to KRY, but with a dataset of U.S. stocks from 2013 to 2021 and with both an environmental score and a greenhouse gas (GHG) emissions intensity score. Their log-emissions intensity is the logarithm of firms' Scope 1 GHG emissions to revenue ratios. The log-emissions intensity has an average cross-sectional correlation of 58% with the MSCI environmental score, so they define their environmental score as the residuals of this score after regression on the log-emissions intensity. The supply coefficients for these metrics are similar in magnitude to the S-score and KRY E-score values reported above.

Note that my estimation (and BCJ's) takes place on five-year windows while KRY and NOS use quarterly data. Intuitively this would be more likely to be an issue if I were comparing price elasticity estimates from these different horizons as both supply and demand are likely to appear more elastic at longer horizons (Gabaix and Koijen, 2021; Van der Beck, 2022). However, it is not clear that there should be a similar horizon-dependence problem for demand coefficients and the similarity of the results from the different sources in Table VI offers some comfort in this regard.

5 Empirical Applications

This section demonstrates how contribution multipliers create a substantial wedge between the impact frontier and frontiers based on *raw externality scores* (illustrated here by ESG ratings), such as the ESG-efficient frontier. In climate and nature finance applications, this wedge matters because the relevant object investor impact on an externality, not the portfolio score itself. Using

investors and brokers have lower average supply coefficients, but of the same sign and order of magnitude, ranging from 1% to 4%. Hedge funds, small active investors and private banking have negative coefficients, ranging from -1% to -4%.

Source	Metric	Supply Coefficient	$R^{(I)}_{1SD}$
BCJ	S score	0.025	0.9%
BCJ	S score	0.126	4.5%
KRY	E score	0.05	0.9%
KRY	E score	0.25	4.3%
NOS	log GHG	-0.023	-0.4%
NOS	log GHG	-0.115	-2.0%
NOS	Residual E	0.031	0.5%
NOS	Residual E	0.155	2.6%

Table VI. **Observed supply shifts for raw externality-score proxies**. This table reports observed supply shifts associated with commonly used score-based metrics from different studies (BCJ, KRY, and NOS), used here as illustrative *raw externality-score proxies*. The original estimates are reported alongside—to facilitate comparison between the studies—implied 'impact returns' $R^{(I)}_{1SD}$ (as in equation 17), which state the change in an investor's required return for an average firm given a one standard deviation change in the underlying metric. The results are shown both for the sample-average coefficients and values five times larger to illustrate the range of supply shifts observed in studies like KRY.

the estimated model parameters from Section 3, I first examine the distribution of contribution multipliers across firms and their relationship with other firm characteristics. I then construct impact frontiers and compare them to frontiers that ignore contribution multipliers. The results reveal that optimizing directly for raw scores without accounting for contribution can lead to portfolios that generate little or even negative expected impact.

5.1 Additionality Multipliers

While the whole contribution multiplier matrix drives the impact frontier, a natural place to first examine the matrix is to understand the multipliers on the diagonal, which I refer to as the additionality multipliers. Figure 3 presents the distribution of additionality multipliers estimated based on the calibrated models from Tables III and IV. Figure 4 then visualizes the relationship between a central estimate of the multipliers and other firm characteristics.

The substantial cross-sectional variation in the additionality multipliers challenges the conventional wisdom that public market investors have limited ability to generate impact, at least at the five year horizons examined here. Some firms indeed have low multipliers (below 10%) and all the more so for the results based on IV. However, some firms have much higher multipliers, suggesting significant heterogeneity in impact potential across firms. This variation is not strongly correlated with traditional firm characteristics, underscoring the need for explicit estimation of contribution multipliers rather than relying on simple proxies. In particular, the additionality multipliers have close to zero correlation with the raw externality-score proxies (illustrated here by ESG scores) for a variety of specifications.

5.2 Impact frontiers

To construct the impact and ESG-efficient frontiers, I use the social score (from BCJ) as a representative *raw externality-score proxy*. To define expected financial performance, I use the value of the investor's financial utility from Equation 2, per unit portfolio value, as the risk-adjusted excess return (i.e., a certainty-equivalent return). This results in more robust and interpretable optimal portfolios than defining the frontier based on nonlinear metrics like portfolio Sharpe ratios.

Figure 5 illustrates the fundamental difference between the impact frontier (solid blue line) and the other frontiers for the baseline model of column (1) in Table III. While the former represents optimal portfolios accounting for contribution multipliers, the ESG-efficient frontier (dashed red line) represents portfolios that optimize directly on ESG scores without accounting for the heterogeneity in contribution. The 'simple' and 'partial' impact frontiers are in the middle as they only ignore off-diagonal terms in the construction of the contribution multiplier matrix. The cross-sectional heterogeneity in the elements of the contribution multiplier matrix, both on and off the diagonal, drives a wedge between each of the frontiers. Each frontier was calculated without restrictions on short selling (the results with such restrictions are qualitatively similar, though with much lower values of investor impact).

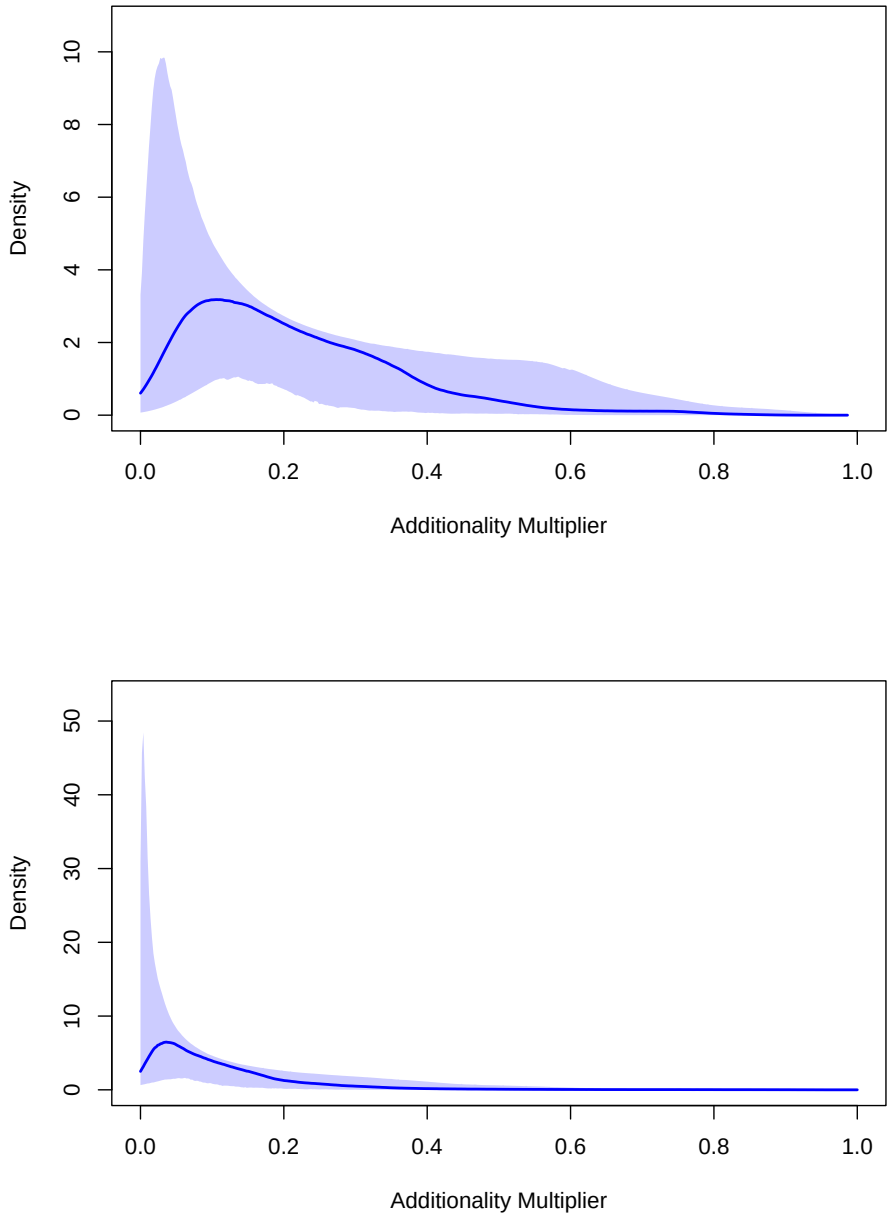


Figure 3. **Distribution of Additionality Multipliers**. The contribution multiplier matrix quantifies the proportion of an investor's supply shift that translates into a change in installed firm capital. The upper panel presents the distribution of the additionality multipliers (the diagonal elements of the contribution multiplier matrix) derived from the results of column (1) in Table III, while the lower panel is based on column (1) in Table IV. The blue shaded areas are the 95% range for 1,000 draws of the contribution multiplier matrix from the estimated coefficient distribution, while the solid blue lines are the median density across these draws.

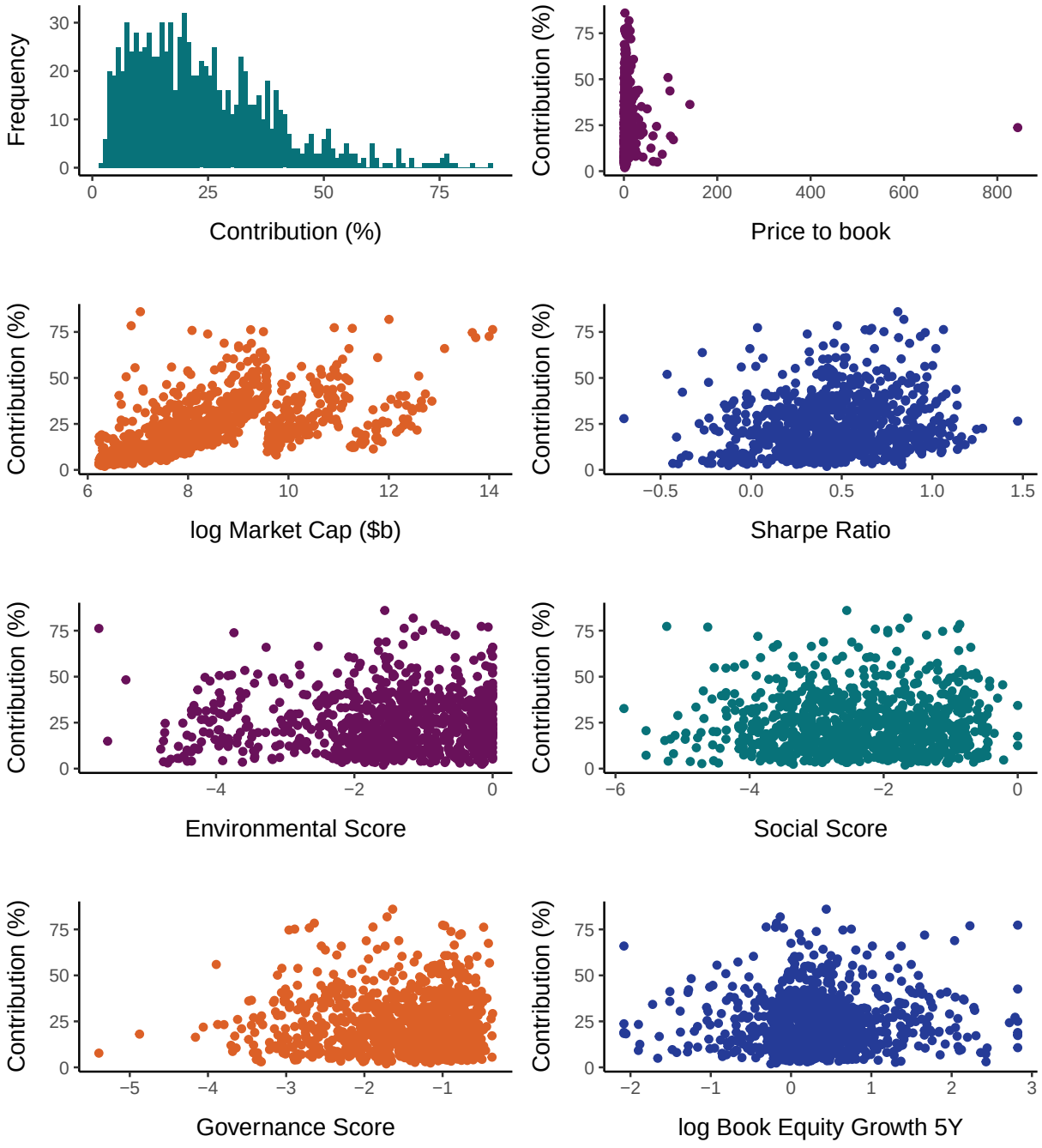


Figure 4. **Additionality Multipliers and Other Firm Characteristics**. The contribution multiplier matrix quantifies the proportion of an investor's supply shift that translates into a change in installed firm capital. The panel in the top-left presents the cross-sectional distribution of the additionality multipliers (the diagonal of the contribution multiplier matrix) in the model based on the parameters of column (1) in Table III. The other panels present scatter plots of the multipliers' relationship with other firm characteristics, showing that additionality multipliers vary substantially across firms but are not strongly correlated with other firm metrics.

Figure 6 presents the frontiers for the model of column (1) in Table IV, which introduces off-diagonal terms into the supply side multiplier matrix ($\boldsymbol{\mathcal{M}}_D$) because of the dependence on industry book equity growth. The key difference this generates is that the frontiers that ignore these terms (the ESG-efficient frontier and the 'simple' impact frontier) actually correspond to negative impact on average. Figure 7 returns to the baseline model to present frontiers for different ESG metrics—the results demonstrate how the wedge between the frontiers depends on the distribution of the underlying firm impact productivity $\boldsymbol{g}$.

The figures show that the cross-sectional heterogeneity in the contribution multipliers drives a wedge between the impact frontier and the ESG-efficient frontier. While the impact frontier shows the best possible return for a given level of investor impact, the ESG-efficient frontier represents portfolios that optimize for ESG scores without accounting for the heterogeneity in contribution. Because the ESG-frontier is not optimizing for impact, but rather portfolio ESG scores, it does not have to curve in the same direction as the impact frontier and indeed in both figures it does not.

The supply-shift coefficients in Table VI show that investors already exhibit meaningful score-based tilts (illustrated here by ESG-related metrics). Comparing to the frontier charts highlights that investors with average score-based tilts are operating in a region where the difference between the impact frontier and the ESG-efficient frontier is economically significant and growing with the strength of those preferences.

Note that the difference between the frontiers is driven purely by the heterogeneity of *contribution multipliers* in the model. If there was no heterogeneity then the impact frontier and ESG-efficient frontier would overlap. This is because if the contribution multipliers are the same for all firms then not adjusting for contribution would merely change your demand coefficient and hence your chosen position on the frontier–it wouldn't change the shape of the frontier. What makes the impact frontier different from the ESG-efficient frontier is that different firms present different opportunities for investor impact per dollar and the impact-optimal policies take this into account.

The fact that the ESG-efficient frontier can produce negative impact is particularly noteworthy. This counterintuitive result occurs because optimizing directly for raw scores (illustrated by ESG

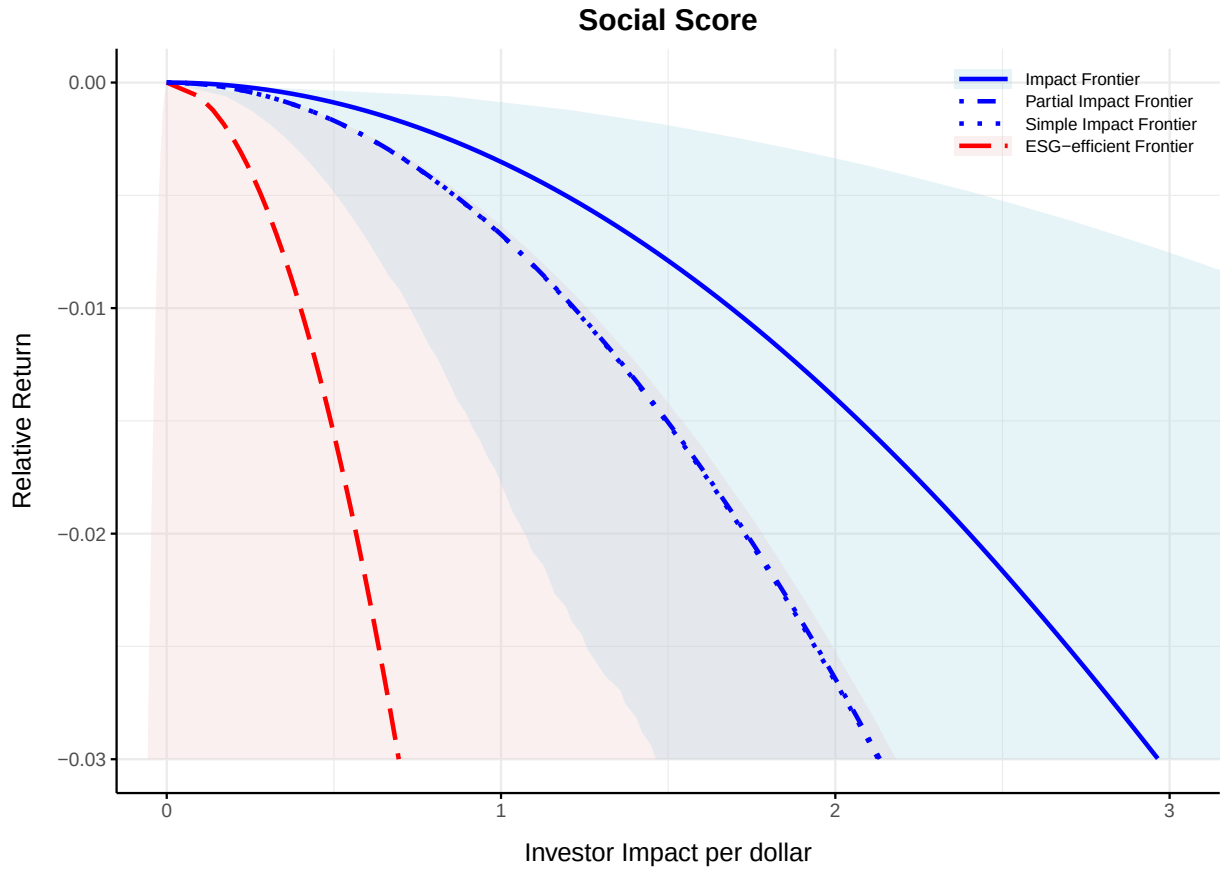


Figure 5. **Impact Frontier for the Social Score**. This figure illustrates different frontiers when using the social score (from BCJ) as an illustrative raw externality-score proxy in the model derived from the results of column (1) in Table III. The solid blue 'Impact Frontier' represents the highest attainable risk-adjusted excess return for each level of investor impact when optimizing using the full contribution multiplier matrix. The dashed red 'ESG-efficient Frontier' represents optimization ignoring the contribution multiplier matrix (though actual investor impact is always calculated using the full matrix). The dotted blue 'Simple Impact Frontier' is based on a version of the contribution multiplier matrix calculated based only on the diagonals of the component matrices, thereby ignoring interfirm systemic effects. The dot dashed blue 'Partial Impact Frontier' is based on only the diagonal of the full contribution multiplier matrix, thereby capturing some but not all systemic effects. The results are presented relative to the baseline portfolio with no impact or holdings-based tilt. Investor impact per dollar is measured on the externality calculated as the aggregate sum of firm social scores and book values and is per dollar of portfolio value shifted. A value of one is equivalent to shifting one dollar of book value from one firm to another firm with a score that is one unit higher. The significant gap between these frontiers demonstrates that ignoring contribution multipliers can substantially reduce or even negate the impact achieved per unit of financial sacrifice. The blue and red shaded areas are the 95% range for 1,000 draws from the estimated parameter distribution for the impact and ESG-efficient frontiers, respectively, while the solid lines are the average for each frontier.

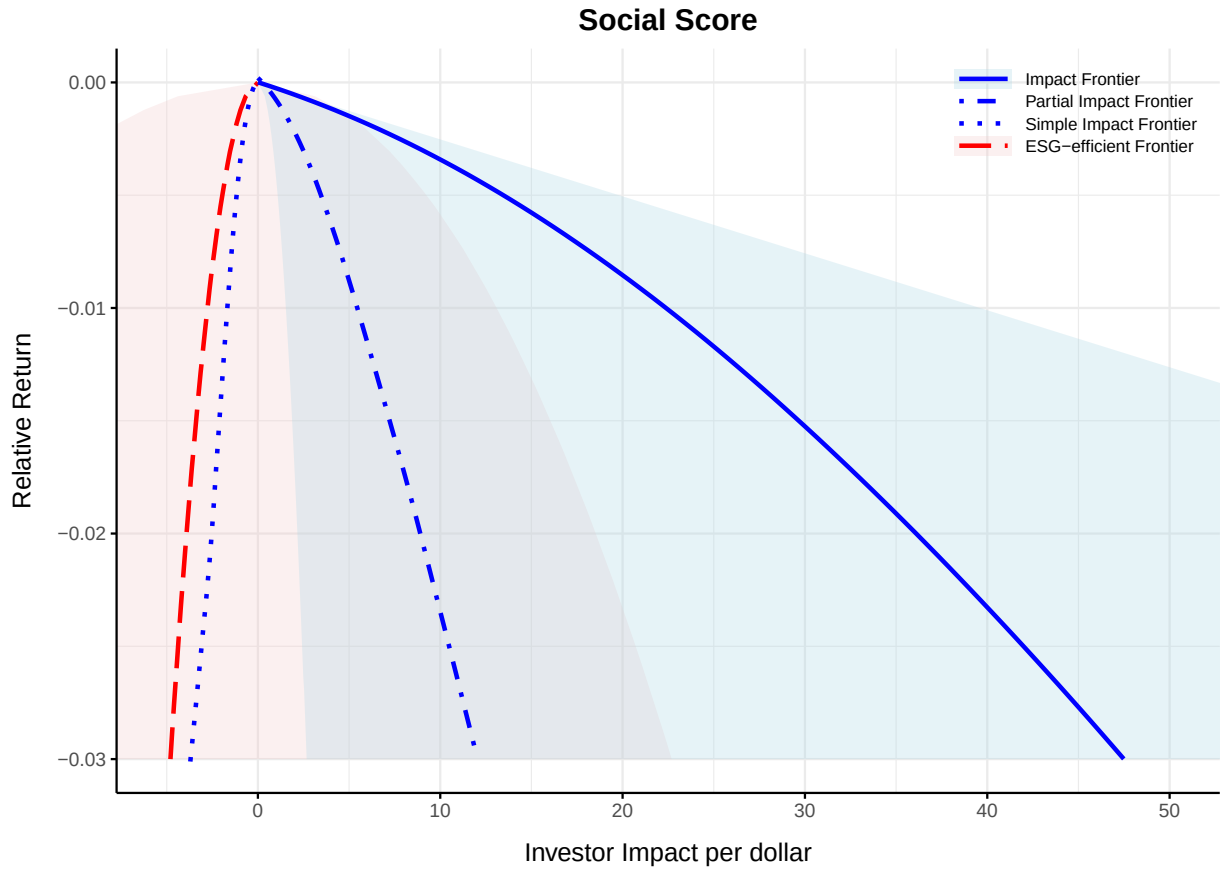


Figure 6. **Impact Frontier for the Social Score with Supply-side Industry Effects**. This figure illustrates different frontiers for the social score in the model derived from the results of column (1) in Table IV. The solid blue 'Impact Frontier' represents the highest attainable risk-adjusted excess return for each level of investor impact when optimizing using the full contribution multiplier matrix. The dashed red 'ESG-efficient Frontier' represents optimization ignoring the contribution multiplier matrix, though actual investor impact is always calculated using the full matrix. The dotted blue 'Simple Impact Frontier' is based on a version of the contribution multiplier matrix calculated based only on the diagonals of the component matrices, thereby ignoring interfirm systemic effects. The dot dashed blue 'Partial Impact Frontier' is based on only the diagonal of the full contribution multiplier matrix, thereby capturing some but not all systemic effects. The results are presented relative to the baseline portfolio with no impact or holdings-based tilt. Investor impact per dollar is measured on the externality calculated as the aggregate sum of firm social scores and book values and is per dollar of portfolio value shifted. A value of one is equivalent to shifting one dollar of book value from one firm to another firm with a score that is one unit higher. The blue and red shaded areas are the 95% range for 1,000 draws from the estimated parameter distribution for the impact and ESG-efficient frontiers, respectively, while the solid lines are the average for each frontier.

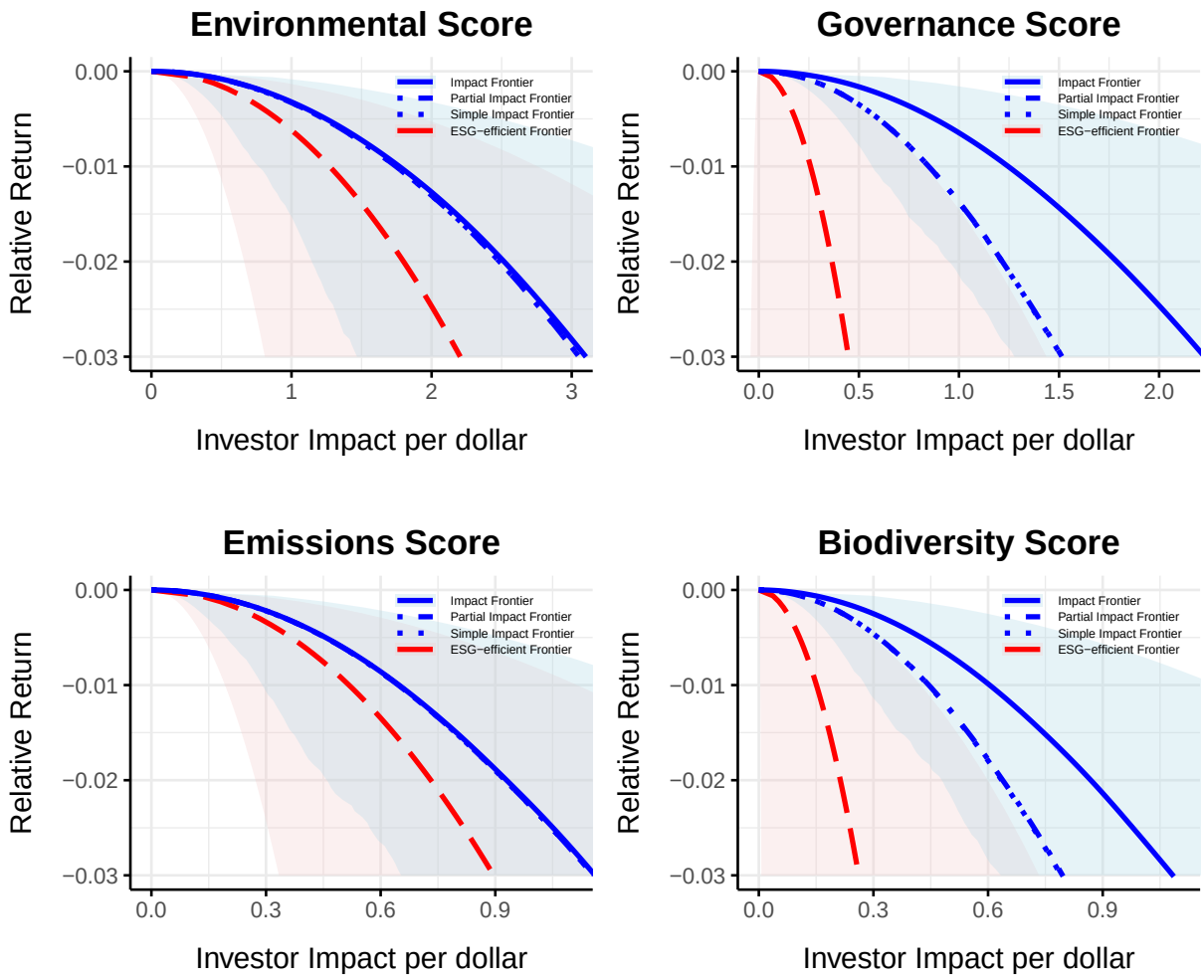


Figure 7. **Impact Frontiers for other metrics**. This figure illustrates frontiers for the environmental, governance, emissions and biodiversity scores in the model derived from the results of column (1) in Table III. The solid blue 'Impact Frontier' represents the highest attainable risk-adjusted excess return for each level of investor impact when optimizing using the full contribution multiplier matrix. The dashed red 'ESG-efficient Frontier' represents optimization ignoring the contribution multiplier matrix (though actual investor impact is always calculated using the full matrix). The dotted blue 'Simple Impact Frontier' is based on a version of the contribution multiplier matrix calculated based only on the diagonals of the component matrices, thereby ignoring interfirm systemic effects. The dot dashed blue 'Partial Impact Frontier' is based on only the diagonal of the full contribution multiplier matrix, thereby capturing some but not all systemic effects. The results are presented relative to the baseline portfolio with no impact or holdings-based tilt. Investor impact per dollar is measured on the externality calculated as the aggregate sum of firm scores and book values and is per dollar of portfolio value shifted. A value of one is equivalent to shifting one dollar of book value from one firm to another firm with a score that is one unit higher. The blue and red shaded areas are the 95% range for 1,000 draws from the estimated parameter distribution for the impact and ESG-efficient frontiers, respectively, while the solid lines are the average for each frontier.

scores) can lead to tilting toward firms with high scores but low contribution multipliers. In such cases, the capital reallocation may shift resources away from firms where marginal investment would generate a larger positive effect on the externality, despite their lower raw scores. This finding highlights the importance of incorporating contribution multipliers when constructing portfolios intended to generate real-world impact.

6 Limitations and Potential Extensions

The model has several key features relative to conventional asset pricing models: the externality and heterogeneous supply and demand elasticities unlock contribution multipliers and investor impact; introducing a price of impact then traces out the impact frontier. While the current framework makes several simplifying assumptions, the core insights about contribution multipliers and their role in investing are likely to be robust to various model extensions. Yet, there are numerous natural refinements and extensions that could be important.

First, it is a single period model, and the microfounded case is for an all-equity economy. Throughout, the single period can be interpreted as a reduced-form representation of portfolio tilts whose real effects materialize over multi-year horizons, consistent with the empirical design.

BCJ include debt in their model and, in their appendix, show how to extend the model to multiple periods. As such, I don't expect such extensions to be problematic. That is not to say such extensions would not be valuable. In particular, I expect that the price of impact could be made endogenous in a multi-period model, with investors balancing financial returns and managing the externality, which will also affect their future welfare.

Other features that could be more explicitly studied within this framework include: the effect of transaction costs, locally increasing returns to scale for selected firms, the 'valley of death' phenomena in some markets (i.e., as a lack of investor supply and an opportunity for investor impact), firm manager preferences beyond shareholder value (e.g., for the externality, cash flows), investor utility depending on the wealth of other impact-focused investors (as collaborators), heterogeneous

and endogenous firm profitability and impact productivity (e.g., as in Pástor, Stambaugh, and Taylor (2021) and Hartzmark and Shue (2022)).

7 Conclusion

By extending the supply and demand system of BCJ to incorporate an externality while maintaining analytical tractability, this paper demonstrates how cross-sectional variation in supply and demand elasticities drives a wedge between the impact frontier and frontiers constructed from raw score tilts (e.g., the ESG-efficient frontier). This result highlights both the potential for meaningful investor impacts in public markets over multi-year horizons and the risk that portfolios tilted only on raw scores may generate little or even negative expected impact if contribution is ignored. The framework is designed to speak directly to climate and nature finance questions in which real-economy effects arise through changes in firms' installed capital.

Asset returns, the impact of investor supply shifts, and the optimal portfolio choice of consequentialist investors all hinge on the aggregate supply and demand elasticities through the contribution multiplier matrix. The off-diagonal elements of this matrix capture systemic effects where supply shifts for one firm affect the equilibrium capital of other firms—effects that are entirely missed by simplified approaches that do not assess impact or assume independent firm responses. Assets characterized by higher supply elasticities and lower demand elasticities offer greater contribution multipliers, making them more attractive opportunities for investor impact and, potentially, for expected returns as the heterogeneous supply side means that non-market risks are priced.

The approach provided here offers a systematic way to address a central question about the role of finance in society: how can investors most effectively allocate capital given the real-world externalities of firms and financial markets? Future research can further refine the estimation of contribution multipliers and explore how they vary across different market segments, time periods, and investor types. Advancing this direction of research offers a substantial opportunity for empirical and theoretical contributions.

References

Allcott, H., G. Montanari, B. Ozaltun, and B. Tan. 2022. An economic view of corporate social impact. *Working Paper* .

Bai, H., K. Hou, H. Kung, E. X. Li, and L. Zhang. 2019. The CAPM strikes back? An equilibrium model with disasters. *Journal of Financial Economics* 131:269–98.

Bai, H., and L. Zhang. 2022. Searching for the equity premium. *Journal of Financial Economics* 143:897–926.

Barber, B., A. Morse, and A. Yasuda. 2021. Impact investing. *Journal of Financial Economics* 139:162–85.

Bauer, R., T. Ruof, and P. Smeets. 2021. Get real! Individuals prefer more sustainable investments. *Review of Financial Studies* 34:3976–4043.

Belo, F., C. Xue, and L. Zhang. 2013. A supply approach to valuation. *The Review of Financial Studies* 26:3029–67.

Berg, F., F. Heeb, and J. F. Kölbel. 2023. The economic impact of ESG ratings. *Working Paper* .

Betermier, S., L. E. Calvet, and E. Jo. 2025. A supply and demand approach to capital markets. *Proceedings of Paris December 2021 Finance Meeting EUROFIDAI-ESSEC* .

Brest, P., R. J. Gilson, and M. A. Wolfson. 2018. How investors can (and can't) create social value. *Journal of Corporation Law* 44:205–32.

Briere, M., and S. Ramelli. 2022. Green sentiment, stock returns, and corporate behavior. *Working Paper* .

Choi, J., X. Tian, Y. Wu, and M. Kargar. 2022. Investor demand, firm investment, and capital misallocation. *Working Paper* .

Dello Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze. 2022. Cross-sectional asset pricing with unsystematic risk. *Working Paper* .

Edmans, A., D. Levit, and J. Schneemeier. 2022. Socially responsible divestment. *Working Paper* .

Favilukis, J., L. Garlappi, and R. Uppal. 2023. Evaluating the impact of portfolio mandates. *Working Paper* .

Gabaix, X., and R. S. J. Koijen. 2021. In search of the origins of financial fluctuations: The inelastic markets hypothesis. *Working Paper* .

Gonçalves, A. S., C. Xue, and L. Zhang. 2020. Aggregation, capital heterogeneity, and the investment CAPM. *The Review of Financial Studies* 33:2728–71.

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. *Working Paper* .

Hartzmark, S. M., and K. Shue. 2022. Counterproductive sustainable investing: The impact elasticity of brown and green firms. *Working Paper* .

Hartzmark, S. M., and A. B. Sussman. 2019. Do investors value sustainability? A natural experiment examining ranking and fund flows. *Journal of Finance* 74:2789–837.

Heeb, F., J. F. Kölbel, F. Paetzold, and S. Zeisberger. 2022. Do investors care about impact? *The Review of Financial Studies* .

Koijen, R. S. J., R. Richmond, and M. Yogo. 2024. Which investors matter for equity valuations and expected returns? *The Review of Economic Studies* 91:2387–424.

Koijen, R. S. J., and M. Yogo. 2019. A demand system approach to asset pricing. *Journal of Political Economy* 127:1475–515.

Kölbel, J. F., F. Heeb, F. Paetzold, and T. Busch. 2020. Can sustainable investing save the world? Reviewing the mechanisms of investor impact. *Organization and Environment* 33:554–74.

Krueger, P., Z. Sautner, and L. T. Starks. 2020. The importance of climate risks for institutional investors. *The Review of Financial Studies* 33:1067–111.

Ma, Y. 2019. Nonfinancial firms as cross-market arbitrageurs. *Journal of Finance* 74:3041–87.

Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42:483–510.

Noh, D., S. Oh, and J. Song. 2023. Unpacking the demand for sustainable equity investing. *Working Paper* .

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2021. Sustainable investing in equilibrium. *Journal of Financial Economics* 142:550–71.

Pavlova, A., and T. Sikorskaya. 2023. Benchmarking intensity. *The Review of Financial Studies* 36:859–903.

Pedersen, L. H., S. Fitzgibbons, and L. Pomorski. 2021. Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics* 550–71.

Riedl, A., and P. Smeets. 2017. Why do investors hold socially responsible mutual funds? *Journal of Finance* 72:2505–50.

Sharpe, W. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19:425–42.

Siemroth, C., and L. Hornuf. 2021. Do retail investors value environmental impact?: A lab-in-the-field experiment with crowdfunders. *Working Paper* .

Starks, L. 2023. Presidential address: Sustainable finance and E, S, & G issues - values versus value. *Journal of Finance* Forthcoming.

Van der Beck, P. 2021. Flow-driven ESG returns. *Swiss Finance Institute Research Paper* .

———. 2022. On the estimation of demand-based asset pricing models. *Swiss Finance Institute Research Paper* .

Yasuda, A. 2023. Impact investing. *Palgrave Encyclopedia of Private Equity* Forthcoming.

Zhang, L. 2017. The investment CAPM. *European Financial Management* 23:545–603.

A General Supply and Demand System

Each active investor's supply curve $\boldsymbol{K}_i(\boldsymbol{P}(1+r_f))$ is a vector-valued function of current asset prices $\boldsymbol{P}$ and the risk-free rate r_f. Aggregate supply is $\boldsymbol{K}_S = (\boldsymbol{I} + \boldsymbol{\kappa}_{\mathcal{P}}) \sum_{i \in \mathcal{A}} \boldsymbol{K}_i$. Similarly, each firm's demand curve may be stacked into the vector-valued function $\boldsymbol{K}_D(\boldsymbol{P})$.

Each investor's supply has an associated *absolute* price elasticity matrix, $\boldsymbol{\zeta}_i$, with elements $\zeta_i(n,m) \equiv -\frac{\partial K_i(n)}{\partial P(m)}$. The aggregate absolute price elasticity matrix for investor supply is $\boldsymbol{\zeta}_S = \sum_{i=1}^{I} \boldsymbol{\zeta}_i$. Similarly, but with the opposite sign convention, the absolute price elasticity matrix for firm demand for capital (i.e., firms' net supply of securities) $\boldsymbol{\zeta}_D$ has elements $\zeta_D(n,m) \equiv \frac{\partial K_D(n)}{\partial P(m)}$. It will be useful to work with 'slope' matrices, $\boldsymbol{\mathcal{M}}_D = \boldsymbol{\zeta}_D^{-1}$ and $\boldsymbol{\mathcal{M}}_S = \boldsymbol{\zeta}_S^{-1}$.

Consider a shock $\boldsymbol{\delta}_i$ to investor i's supply curve. The market clearing conditions, for each firm asset and to ensure the risk-free asset is in zero net supply, form the system:

$$\boldsymbol{F}(\boldsymbol{P}, r_f, \boldsymbol{\delta}_i) = \begin{bmatrix} \boldsymbol{K}_D(\boldsymbol{P}) - \boldsymbol{K}_S(\boldsymbol{P}, r_f) - (\boldsymbol{I} + \boldsymbol{\kappa}_{\mathcal{P}})\boldsymbol{\delta}_i \\ \boldsymbol{P}'(\boldsymbol{K}_D(\boldsymbol{P}) - \boldsymbol{K}_0) - \sum_i \mathrm{C}_i \end{bmatrix} = \boldsymbol{0} \tag{22}$$

where $\sum_i \mathrm{C}_i$ is the initial wealth in cash of the investors.

When the market clears $\boldsymbol{K} = \boldsymbol{K}_D$ and denoting $\boldsymbol{B} \equiv \partial_{r_f} \boldsymbol{K}_S = \frac{1}{1+r_f}(\nabla_{\boldsymbol{P}} \boldsymbol{K}_S)\boldsymbol{P} = \frac{1}{1+r_f} \boldsymbol{\mathcal{M}}_S^{-1} \boldsymbol{P}$, the Jacobian is:

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{\zeta}_D + \boldsymbol{\zeta}_S & \boldsymbol{B} \\ (\boldsymbol{K} - \boldsymbol{K}_0)' + \boldsymbol{P}'\boldsymbol{\zeta}_D & \boldsymbol{0} \end{bmatrix} \tag{23}$$

By the implicit function theorem:

$$\begin{bmatrix} \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\delta}_i} \\ \frac{\partial r_f}{\partial \boldsymbol{\delta}_i} \end{bmatrix} = -\boldsymbol{J}^{-1} \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{\delta}_i} = -\boldsymbol{J}^{-1} \begin{bmatrix} -(\boldsymbol{I} + \boldsymbol{\kappa}_{\mathcal{P}}) \\ \boldsymbol{0}' \end{bmatrix} \tag{24}$$

Using the block matrix inverse formula, the identity $\left(\boldsymbol{\mathcal{M}}_D^{-1} + \boldsymbol{\mathcal{M}}_S^{-1}\right)^{-1} = \boldsymbol{\mathcal{M}}_D\big(\boldsymbol{\mathcal{M}}_D + \boldsymbol{\mathcal{M}}_S\big)^{-1}\boldsymbol{\mathcal{M}}_S$, and defining $\boldsymbol{D} \equiv ((\boldsymbol{K} - \boldsymbol{K}_0)'\boldsymbol{\mathcal{M}}_D + \boldsymbol{P}')(\boldsymbol{\mathcal{M}}_D + \boldsymbol{\mathcal{M}}_S)^{-1}\boldsymbol{\mathcal{M}}_S$ and $\boldsymbol{\mathcal{C}} \equiv (\boldsymbol{\mathcal{M}}_D + \boldsymbol{\mathcal{M}}_S)^{-1}\boldsymbol{\mathcal{M}}_S\left(\boldsymbol{I}_N - \frac{\boldsymbol{BD}}{\boldsymbol{DB}}\right)(\boldsymbol{I} + \boldsymbol{\kappa}_{\mathcal{P}})$, the relevant derivatives with respect to $\boldsymbol{\delta}_i$ are:

$$\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{\delta}_i} = \boldsymbol{\mathcal{M}}_D\boldsymbol{\mathcal{C}} \tag{25}$$

$$\frac{\partial r_f}{\partial \boldsymbol{\delta}_i} = \frac{\boldsymbol{D}}{\boldsymbol{DB}}(\boldsymbol{I} + \boldsymbol{\kappa}_{\mathcal{P}}) \tag{26}$$

$$\frac{\partial \boldsymbol{K}}{\partial \boldsymbol{\delta}_i} = \frac{\partial \boldsymbol{K}_D}{\partial \boldsymbol{\delta}_i} = \boldsymbol{\mathcal{C}} \tag{27}$$

$$\frac{\partial \boldsymbol{K}_i}{\partial \boldsymbol{\delta}_i} = \boldsymbol{I} - \boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{M}}_D\boldsymbol{\mathcal{C}} \tag{28}$$

Each investor has mean-variance utility plus an externality term:

$$U_i = \boldsymbol{a}_i'\boldsymbol{K}_i - (1 + r_f)\boldsymbol{P}'\left(\boldsymbol{K}_i - \boldsymbol{K}_{i,0} - \operatorname{diag}\left(\frac{\boldsymbol{K}_{i,0}}{\boldsymbol{K_0}}\right)(\boldsymbol{K} - \boldsymbol{K}_0)\right) - \frac{\gamma_i}{2W_i}\boldsymbol{K}_i'\boldsymbol{\Sigma}_{CF,i}\boldsymbol{K}_i + \gamma_i^g\boldsymbol{g}'\boldsymbol{K}$$

In the infinitesimal limit, where $\boldsymbol{K}_i \to \boldsymbol{0}$ and $\boldsymbol{K}_{i,0} \to \boldsymbol{0}$ but $\frac{\boldsymbol{K}_i}{W_i} \neq \boldsymbol{0}$, then $\frac{\partial \boldsymbol{K}_i}{\partial \boldsymbol{\delta}_i} \approx \boldsymbol{I}$ and the first-order condition for optimal $\boldsymbol{\delta}_i$ is:

$$\frac{\partial U_i}{\partial \boldsymbol{\delta}_i} = \boldsymbol{a}_i' - (1 + r_f)\boldsymbol{P}' - \frac{\gamma_i}{W_i}(\boldsymbol{\Sigma}_{CF,i}\boldsymbol{K}_i)' + \gamma_i^g\boldsymbol{g}'\frac{\partial \boldsymbol{K}}{\partial \boldsymbol{\delta}_i} = \boldsymbol{0}$$

where any dependence of $W_i = \boldsymbol{P}'\boldsymbol{K}_{i,0} + C_i$ on $\boldsymbol{\delta}_i$ leads only to higher-order terms. So, the optimal supply curve for investor i is simply:

$$\boldsymbol{I}_i\boldsymbol{P} = \boldsymbol{\mu}_i - \boldsymbol{\mathcal{M}}_i\boldsymbol{K}_i \tag{29}$$

with $\boldsymbol{\mu}_i = \frac{1}{1+r_f}\boldsymbol{I}_i(\boldsymbol{a}_i + \gamma_i^g\boldsymbol{\mathcal{C}}'\boldsymbol{g})$ and $\boldsymbol{\mathcal{M}}_i = \frac{1}{1+r_f}\frac{\gamma_i}{W_i}\boldsymbol{I}_i\boldsymbol{\Sigma}_{CF,i}\boldsymbol{I}_i'$.

A.1 Investor impact

The inverse of $\boldsymbol{\mathcal{M}}_T = \boldsymbol{M}_D + \boldsymbol{M}_S$ may be decomposed as follows:

$$\boldsymbol{\mathcal{M}}_T^{-1} = \begin{pmatrix} 1 \\ -\boldsymbol{\beta}_{T,n} \end{pmatrix} (\boldsymbol{\mathcal{M}}_T(n,n) - \boldsymbol{\mathcal{M}}_T(n,-n)\boldsymbol{\beta}_{T,n})^{-1} \begin{pmatrix} 1 \\ -\tilde{\boldsymbol{\beta}}_{T,n} \end{pmatrix}' + \begin{pmatrix} 0 & 0 \\ 0 & (\boldsymbol{\mathcal{M}}_T(-n,-n))^{-1} \end{pmatrix},$$

with $\boldsymbol{\beta}_{T,n} = (\boldsymbol{\mathcal{M}}_T(-n,-n))^{-1}\boldsymbol{\mathcal{M}}_T(-n,n)$ and $\tilde{\boldsymbol{\beta}}_{T,n} = (\boldsymbol{\mathcal{M}}_T(-n,-n))^{-1}\boldsymbol{\mathcal{M}}_T(n,-n)'$.

The investor impact of a supply shift $\delta\boldsymbol{K_i}$ is, to first order, $II_i = \boldsymbol{g}'\delta\boldsymbol{K} = \boldsymbol{g}'\frac{\partial \boldsymbol{K}}{\partial \boldsymbol{K}_i}\delta\boldsymbol{K_i} = \boldsymbol{g}'\boldsymbol{\mathcal{C}}\delta\boldsymbol{K_i}$, and using the inverse matrix decomposition below, it may be decomposed as:

$$II_i = \sum_{m,n} g(n)\mathcal{C}^E(n)\mathcal{C}(n,m)\delta K_i(m),$$

with $\mathcal{C}^E(n) \equiv \left(1 - \frac{\boldsymbol{g}'(-n)}{g(n)}\boldsymbol{\beta}_{T,n}\right)$ and $\mathcal{C}(n,m) \equiv \frac{\mathcal{M}_S(n,n)}{\mathcal{M}_T(n,n)} \frac{\left(\delta_{m,n} - \frac{\mathcal{M}_S(n,)\boldsymbol{BD}(,m)}{\mathcal{M}_S(n,n)\boldsymbol{DB}}\right)}{\left(1 - \frac{\boldsymbol{\mathcal{M}}_T(n,-n)}{\mathcal{M}_T(n,n)}\boldsymbol{\beta}_{T,n}\right)}(1 + \kappa_{\mathcal{P}}(m))$.

A.2 Aggregate Demand and Covariance Matrix

The individual investor supply curves can be rearranged to be appropriate to aggregate,

$$\boldsymbol{I}_i'\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{I}_i\boldsymbol{P} = \frac{1}{1+r_f}\boldsymbol{I}_i'\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{I}_i\left((\boldsymbol{a}_i + \boldsymbol{h}_i) - \frac{1}{2}(\boldsymbol{\Sigma}_{CF,D}\boldsymbol{K}) + \gamma_i^g\boldsymbol{\mathcal{C}}'\boldsymbol{g}_i\right) - \boldsymbol{K}_i,$$

indicating that the aggregate investor supply slope matrix, obtained by aggregating the above equation and the applying the inverse of the term on the before $\boldsymbol{P}$, will have $\boldsymbol{\mathcal{M}}_S = \left(\sum_{i\in\mathcal{A}} \boldsymbol{I}_i'\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{I}_i\right)^{-1}$.

Focusing on the covariance matrices contained within each $\mathcal{M}$ matrix, for the microfounded case, the investor-level inverse covariance matrix is

$$(\boldsymbol{I}_i\boldsymbol{\Sigma}_{CF,i}\boldsymbol{I}_i')^{-1} = \boldsymbol{I}_i\boldsymbol{\sigma}_{idio,CF}^{-2}\boldsymbol{I}_i' - \boldsymbol{I}_i\boldsymbol{\sigma}_{idio,CF}^{-2}\boldsymbol{b}_{CF}(\boldsymbol{I}_{F_{CF}} + \boldsymbol{b}_{CF}'\boldsymbol{I}_i'\boldsymbol{I}_i\boldsymbol{\sigma}_{idio,CF}^{-2}\boldsymbol{I}_i'\boldsymbol{I}_i\boldsymbol{b}_{CF})^{-1}\boldsymbol{b}_{CF}'\boldsymbol{\sigma}_{idio,CF}^{-2}\boldsymbol{I}_i'.$$

The latter term in the expression above is nonlinear in the investor's opportunity set and is not analytically tractable in general without approximation. However, if (as in Merton (1987)) assets 1 to F_{CF} are forward contracts on each factor and all investors have access to these contracts, with $\sigma^2_{idio,CF}(1:F_{CF})=0$ and $\boldsymbol{b}_{CF}(1:F_{CF},)=\boldsymbol{I}_{F_{CF}}$, then

$$(\boldsymbol{I}_i\boldsymbol{\Sigma}_{CF,i}\boldsymbol{I}_i')^{-1}=\begin{pmatrix}\boldsymbol{I}_{F_{CF}}+\tilde{\boldsymbol{b}}_{CF}'\boldsymbol{I}_i'\boldsymbol{I}_i\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\boldsymbol{I}_i'\boldsymbol{I}_i\tilde{\boldsymbol{b}}_{CF} & -\tilde{\boldsymbol{b}}_{CF}'\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\boldsymbol{I}_i' \\ -\boldsymbol{I}_i\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\tilde{\boldsymbol{b}}_{CF} & \boldsymbol{I}_i\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\boldsymbol{I}_i'\end{pmatrix},$$

with $\tilde{\boldsymbol{b}}_{CF}=\boldsymbol{b}_{CF}(-(1:F_{CF}),)$ and $\tilde{\boldsymbol{\sigma}}_{idio,CF}=\boldsymbol{\sigma}_{idio,CF}(-(1:F_{CF}),-(1:F_{CF}))$, and aggregating:

$$\boldsymbol{\Sigma}_{CF,S}=\begin{pmatrix}\boldsymbol{I}_{F_{CF}}+\tilde{\boldsymbol{b}}_{CF}'\operatorname{diag}(\boldsymbol{q})\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\tilde{\boldsymbol{b}}_{CF} & -\tilde{\boldsymbol{b}}_{CF}'\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\operatorname{diag}(\boldsymbol{q}) \\ -\operatorname{diag}(\boldsymbol{q})\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\tilde{\boldsymbol{b}}_{CF} & \operatorname{diag}(\boldsymbol{q})\tilde{\boldsymbol{\sigma}}^{-2}_{idio,CF}\end{pmatrix}^{-1}=\boldsymbol{b}_{CF}\boldsymbol{b}_{CF}'+\operatorname{diag}(\boldsymbol{q})^{-1}\boldsymbol{\sigma}^2_{idio,CF}.$$

A.3 Asset Returns and the Market

In equilibrium the capital amounts will be $\boldsymbol{K}=\boldsymbol{\mathcal{M}}_T^{-1}\boldsymbol{\mu}_T$, with $\boldsymbol{\mu}_T\equiv\boldsymbol{\mu}_S-\boldsymbol{\mu}_D$, and prices will be $\boldsymbol{P}=\boldsymbol{\mu}_S-\boldsymbol{\mathcal{M}}_S\boldsymbol{\mathcal{M}}_T^{-1}\boldsymbol{\mu}_T=(\boldsymbol{I}-\boldsymbol{\mathcal{M}}_S\boldsymbol{\mathcal{M}}_T^{-1})\boldsymbol{\mu}_S+\boldsymbol{\mathcal{M}}_S\boldsymbol{\mathcal{M}}_T^{-1}\boldsymbol{\mu}_D$. Next, define some useful variables: the market value, $V_M\equiv\boldsymbol{P}'\boldsymbol{K}$; the variance of the market return, $\operatorname{Var}(R_M)\equiv\boldsymbol{K}'\Sigma_{CF}\boldsymbol{K}/V_M^2$; the vector of asset covariances with the market return, $\mathbf{Cov}_M\equiv\operatorname{diag}(\boldsymbol{P})^{-1}\Sigma_{CF}\boldsymbol{K}/V_M$; and the vector of standard market betas, $\boldsymbol{\beta}_M\equiv\mathbf{Cov}_M/\operatorname{Var}(R_M)$.

Rearranging the aggregate demand equation and noting that $\boldsymbol{\Sigma}_{CF}\boldsymbol{K}=V_M\operatorname{Var}(R_M)\operatorname{diag}(\boldsymbol{P})\boldsymbol{\beta}_M$:

$$\boldsymbol{a}-\frac{1}{2}\boldsymbol{\Sigma}_{CF,D}\boldsymbol{K}=(1+r_f)\boldsymbol{P}+\operatorname{diag}(\boldsymbol{P})\Big(\boldsymbol{\lambda}+V_M\operatorname{Var}(R_M)\frac{\gamma}{W_0}\boldsymbol{\beta}_M\Big)$$

with $\boldsymbol{\lambda}\equiv\operatorname{diag}(\boldsymbol{P})^{-1}\left[(1+r_f)\Big(\boldsymbol{\mathcal{M}}_S-\frac{1}{(1+r_f)}\frac{\gamma}{W_0}\boldsymbol{\Sigma}_{CF}\Big)\boldsymbol{K}-\tilde{\boldsymbol{h}}-\tilde{\gamma}^g\boldsymbol{\mathcal{C}}'\tilde{\boldsymbol{g}}\right]$.

The expected excess return on the market is then:

$$E[R_M] - r_f = (\boldsymbol{a} - \frac{1}{2}\boldsymbol{\Sigma}_{CF,D}\boldsymbol{K})'\boldsymbol{K}/V_M - (1+r_f) = \frac{\gamma}{W_0}V_M Var(R_M) + \lambda_M,$$

with $\lambda_M \equiv \boldsymbol{\lambda}'\boldsymbol{V}/V_M$. Similarly, the expected excess returns are:

$$E[\boldsymbol{R}] - r_f = \text{diag}(\boldsymbol{P})^{-1}(\boldsymbol{a} - \frac{1}{2}\boldsymbol{\Sigma}_{CF,D}\boldsymbol{K}) - (1+r_f) = \boldsymbol{\beta}_M(E[R_M] - r_f) + (\boldsymbol{\lambda} - \boldsymbol{\beta}_M\lambda_M).$$

A.4 Impact Frontier

For simplicity, I ignore opportunity set heterogeneity in this subsection. The relevant part of the investor's utility is

$$\begin{aligned} U_i^{(IF)} &\approx (\boldsymbol{a}_i + \boldsymbol{h}_i - (1+r_f)\boldsymbol{P})'\boldsymbol{K}_i - \frac{1+r_f}{2}\boldsymbol{K}_i'\boldsymbol{\mathcal{M}}_i\boldsymbol{K}_i \\ &= \frac{1}{2}(\boldsymbol{a}_i + \boldsymbol{h}_i - (1+r_f)\boldsymbol{P})'\boldsymbol{\mathcal{M}}_i^{-1}(\boldsymbol{a}_i + \boldsymbol{h}_i - (1+r_f)\boldsymbol{P}) - \frac{1}{2}(\delta\boldsymbol{K}_i^g)'\boldsymbol{\mathcal{M}}_i\delta\boldsymbol{K}_i^g \\ &= U_{i,0} - \frac{1}{2}\frac{(\gamma_i^g)^2}{(1+r_f)}\boldsymbol{g}_i'\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{C}}'\boldsymbol{g}_i = U_{i,0} - \frac{1+r_f}{2(\boldsymbol{g}_i'\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{C}}'\boldsymbol{g}_i)}(II_i^g)^2, \end{aligned}$$

The slope of the impact frontier at any tangent point is $\frac{\partial U_i^{(IF)}}{\partial II_i^g} = -\frac{1+r_f}{(\boldsymbol{g}_i'\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{C}}'\boldsymbol{g}_i)}II_i^g = -\gamma_i^g$. The 'inefficient' impact frontier follows similarly and by the Cauchy-Schwarz inequality $\frac{1}{(\boldsymbol{g}_i'\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{C}}'\boldsymbol{g}_i)} \leq \frac{\boldsymbol{g}_i'\boldsymbol{\mathcal{C}}^*\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{C}}^{*\prime}\boldsymbol{g}_i}{(\boldsymbol{g}_i'\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{M}}_i^{-1}\boldsymbol{\mathcal{C}}^{*\prime}\boldsymbol{g}_i)^2}$ so the impact frontier is greater than or equal to the inefficient frontier.