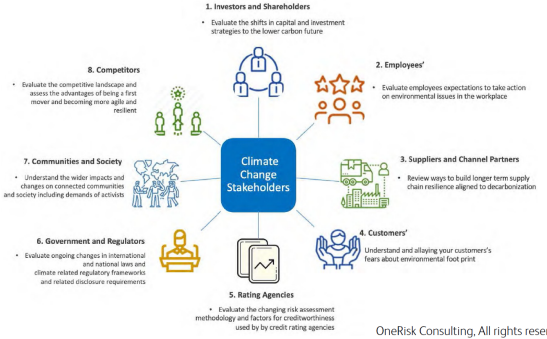
Impact Trickles Down: A General Equilibrium Theory of Stakeholder Exit and Engagement

Briana Chang and Harrison Hong

HKUST and Columbia University

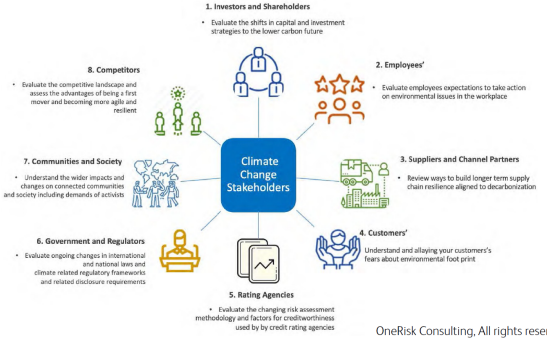
May, 2026

Motivation: Rising pressure from multiple purpose-driven stakeholders



- **Exit** vs. **Engage**? (Hirschman (1970))
  - Engagement dominates exit (Broccardo et al. (2022), Dimson et al. (2015))
  - Empirically, exits seem ineffective, as they exit to firms that are already "good".

Motivation: Rising pressure from multiple purpose-driven stakeholders

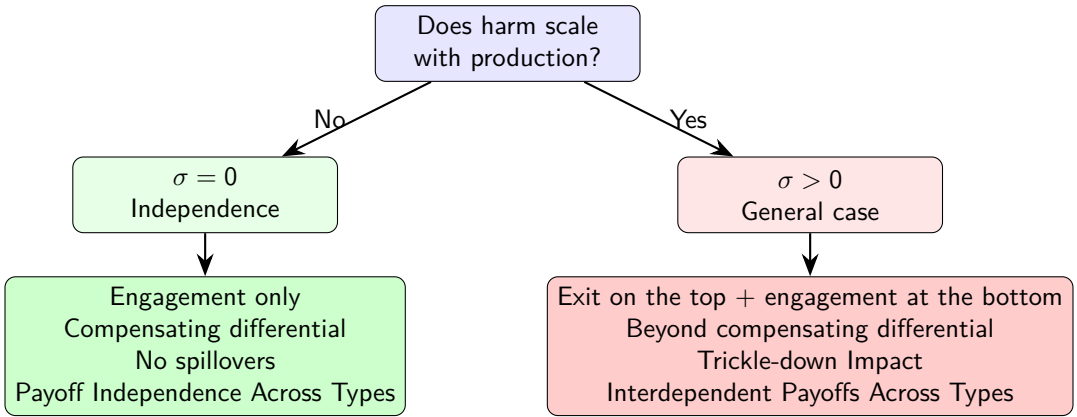


- **Exit** vs. **Engage**? (Hirschman (1970))
  - Engagement dominates exit (Broccardo et al. (2022), Dimson et al. (2015))
  - Empirically, exits seem ineffective, as they exit to firms that are already "good".
  - Yet, exit remains prevalent among large stakeholders

What We Do

- A tractable **GE** model with multiple stakeholder groups (workers, banks, suppliers, . . .) interacting in competitive, multi-sided markets
  - heterogeneous preferences and productivity levels
- Exit vs. engagement as optimal choices to **break or maintain prior relationships**
  - firms: hiring suboptimal stakeholders vs. mitigating
  - purpose-driven stakeholders: joining a different team vs. subsidizing the firm
- Key objects: equilibrium prices and team composition

What We Learned: Determinants of exit vs. engagement



Implications for empirical research

- Empirical studies find that firm-level divestment/exit has **limited impact**
  - Kacperczyk and Luis-Peydro (2022): green banks divest, brown banks substitute
  - Firm-level diff-in-diff: near-zero effects
- Consistent with our model predictions:
  - Large stakeholders exit to purpose-aligned ("clean") firms
  - These firms already mitigate $\rightarrow$ exit looks ineffective at firm level
- **Take-away:** Firm-level studies miss **general equilibrium spillovers**
  - A trickle-down effect through reallocation

Contribution to the Literature

- Exit vs. voice in partial EQ, or single stakeholders
  - SRI/ESG exit affects financing Heinkel et al. (2001), Hong et al. (2021), Pástor et al. (2021), Pedersen et al. (2021), Oehmke and Opp (2023).
  - Voting/voice may dominate in micro settings: Broccardo et al. (2022).
- Two-sided Matching
  - 1-D: Sattinger (1979), Tervio (2008), Gabaix and Landier (2008).
  - 2-D: Hatfield and Kominers (2015), Dupuy and Galichon (2014), Lindenlaub (2017), Chiappori et al. (2016), Chiappori et al. (2018).
- Multi-sided, 1-D: Boerma et al. (2025)
- **This paper:** multi-sided, 2-D - tractable characterization w. iterative algorithm

Roadmap

1. Model
2. Independence Benchmark ($\sigma = 0$)
3. General Case ($\sigma > 0$)
4. Exit vs. Engagement (Application: Bank Decarbonization)

Environment: Agents

- N types of stakeholders (workers, banks, etc) + firms, continuum of each type
  - heterogeneous ability x_ℓ
  - (separable) production function $y(\boldsymbol{x}) = \prod_{\ell=1}^{N+1} x_\ell$.
    - e.g. Cobb-Douglas $AK^\alpha L^{(1-\alpha)}$

- Purpose-driven ($\theta_\ell = 1$) stakeholders experience **disutility from harm**:

$$u(p, e \,|\, \theta_\ell) = p - \theta_\ell \cdot \psi(e)$$

  where $\theta_\ell \in \{0, 1\}$, p = transfers, e = firm's social harm

- Stakeholders are two-dimensional $a_\ell \equiv (x_\ell, \theta_\ell)$ $\forall \ell$
  - i.i.d, where the fraction of purpose-driven stakeholders is λ_ℓ

Environment: Production, Harm, and Mitigation

- Social harm and output:

$$\phi(y) = \phi_0 + \sigma \cdot y$$

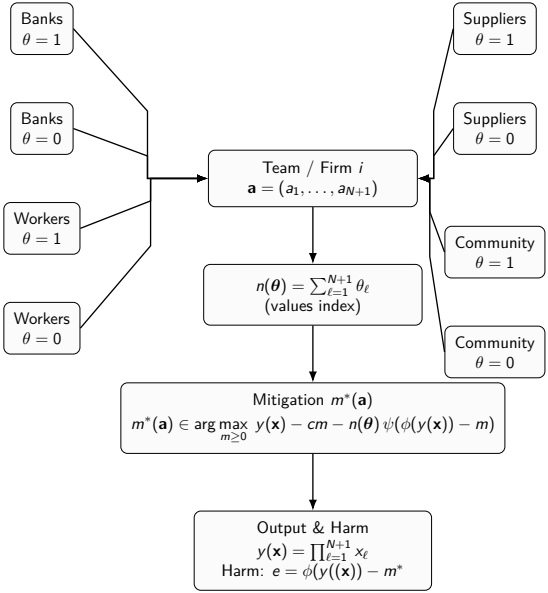
  - $\sigma = 0$: harm is independent with productivity
  - $\sigma > 0$: harm scales with production (pollution, emissions)

- Firms can **mitigate** harm at a linear cost c

- Harm after mitigation:

$$e = \phi(y) - m$$

Schematic of Model



Competitive Equilibrium

A competitive equilibrium consists of:

1. A matching γ between firms and stakeholders
2. A firm mitigation policy $m^*(\boldsymbol{a})$.
3. Utilities $\{U_\ell(a_\ell)\}$ for each type

Subject to:

- Mitigation policy maximizes joint surplus
- No profitable deviation by forming a new team

$$U_\ell(a_\ell) = \max_{\{a_{\ell'}\}_{\ell' \in L \setminus \{\ell\}}} \Lambda(\{a_{\ell'}\}_{\ell' \in L \setminus \{\ell\}}, a_\ell) - \sum_{\ell' \in L \setminus \{\ell\}} U(a_{\ell'}). \tag{1}$$

- Markets clear (marginals preserved).

Existence and efficiency: multi-sided matching with transfers (Boerma et al. (2025))

Roadmap

1. Model
2. Independence Benchmark ($\sigma = 0$)
3. General Case ($\sigma > 0$)
4. Exit vs. Engagement (Application: Bank Decarbonization)

When Harm Doesn't Scale with Production ($\sigma = 0$)

Proposition (Independence Benchmark): When $\sigma = 0$, sorting on productivity is **independent** of preferences

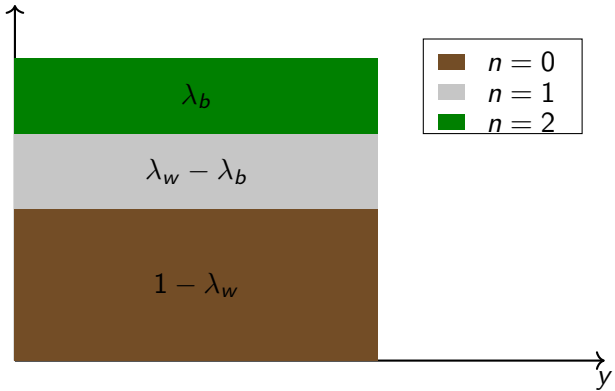
(i) Positive sorting on productivity

(ii) Positive sorting on preferences
- Mitigation is non-rival, purpose-stakeholders cluster to share the costs
- The distribution of the purpose index is determined by clustering

(iii) Compensating differential
- Firm at ranking i is indifferent between $(x_\ell[i], 0)$ and $(x_\ell[i], 1)$

Independence Benchmark: Illustration for Banks and Workers

Figure 1: Distribution of purpose-driven index across y, where $\lambda_w \geq \lambda_b$



Independence: Engagement Only, No Spillovers

Corollary (Engagement Only):

- An increase in λ_ℓ (more purpose-driven stakeholders of type ℓ):
  - Treated stakeholders work for **same-sized firms** as before
  - Only compensating differentials change
  - All adjustment is through **engagement** (voice), not exit

Corollary (No Spillovers):

- The preference shock to type ℓ does **not** affect:
  - Equilibrium allocation of any agent
  - Utilities of stakeholders of other types $\ell' \neq \ell$

$\Rightarrow$ This is why existing literature can study one stakeholder type at a time.

Roadmap

1. Model
2. Independence Benchmark ($\sigma = 0$)
3. General Case ($\sigma > 0$)
4. Exit vs. Engagement (Application: Bank Decarbonization)

When Harm Scales with Production ($\sigma > 0$)

Everything changes.

- Higher output $\rightarrow$ more harm $\rightarrow$ productivity is discounted whenever $n \geq 1$

$$\Omega_y(y, n) = \begin{cases} 1 - c\sigma, & n \geq 1, \\ 1 & n = 0. \end{cases}$$

- Pure profit-driven teams become relatively valuable than its counterparts

When Harm Scales with Production ($\sigma > 0$)

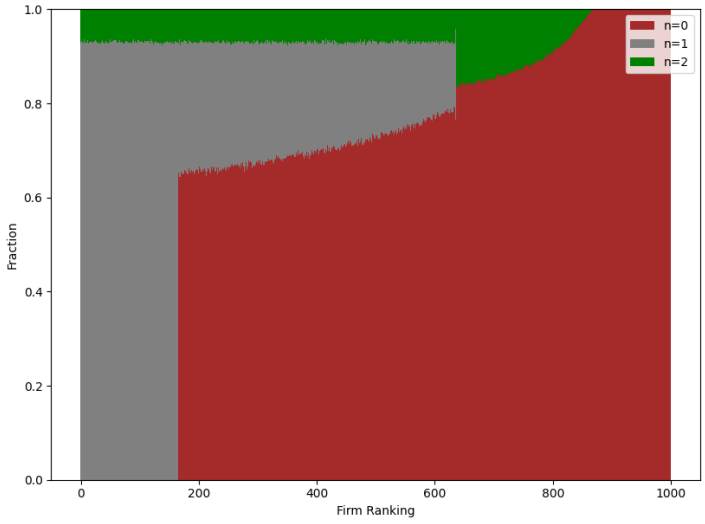
Everything changes.

- Higher output $\rightarrow$ more harm $\rightarrow$ productivity is discounted whenever $n \geq 1$

$$\Omega_y(y, n) = \begin{cases} 1 - c\sigma, & n \geq 1, \\ 1 & n = 0. \end{cases}$$

- Pure profit-driven teams become relatively valuable than its counterparts
- **Result:** High-productivity firms compete for less productive but pure profit-driven teams
  - Profit-driven stakeholders could earn additional **rents**
  - Purpose-driven stakeholders cluster together at **smaller** firms
  - **Full separation at the top** of the productivity distribution

Segmentation at the Top, Mixing at the Bottom



Why Mixing Occurs at the Bottom: avoiding mitigation vs. worse matches

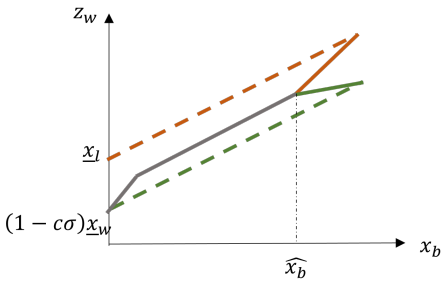


Figure 2: Dashed (vs. solid) lines represent the effective z-index of the matching workers for banks under balanced supply (vs. $\lambda_w > \lambda_b$).

- From the viewpoint of profit-driven banks: effective ranking is $z(x_w, \theta)$
- The productive profit-workers are outbid by the top
- For low profit-driven banks ($x_b < \hat{x}_b$): indifferent between $x_w = (1 - c\sigma)x'_w$

Rents for Certain Profit-Driven Stakeholders

Proposition (Rents): Profit-driven stakeholders in pure profit-driven teams earn rents beyond compensating differentials

- All profit-driven stakeholders of type 1 (since they're scarce)
- For types $\ell \geq 2$: only if above the mixing cutoff: $x_\ell > \hat{x}_\ell$
- Below cutoff: no rent (standard compensating differentials)

- **Contrast with $\sigma = 0$:** No rents under independence—only compensating differentials

Rents for Certain Profit-Driven Stakeholders

Proposition (Rents): Profit-driven stakeholders in pure profit-driven teams earn rents beyond compensating differentials

- All profit-driven stakeholders of type 1 (since they're scarce)
- For types $\ell \geq 2$: only if above the mixing cutoff: $x_\ell > \hat{x}_\ell$
- Below cutoff: no rent (standard compensating differentials)

- **Contrast with $\sigma = 0$:** No rents under independence—only compensating differentials

- **Predictions**: wage premium larger than interest rate premium (as $\lambda_w > \lambda_b$)

Roadmap

1. Model
2. Independence Benchmark ($\sigma = 0$)
3. General Case ($\sigma > 0$)
4. Exit vs. Engagement (Application: Bank Decarbonization)

More Green Banks, More Separation on the top

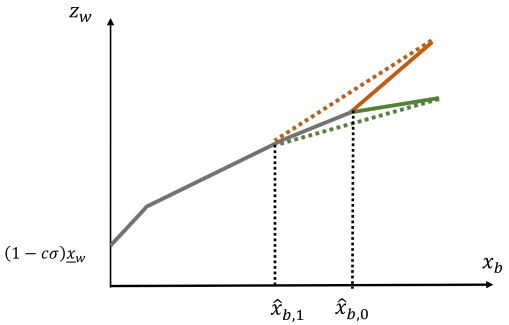


Figure 3: The solid (dotted) line represents the matching before (after) the preference shock that increases the measure of purpose-driven banks.

Application: Post-2015 Paris Agreement (λ_b: 7% $\rightarrow$ 15%)

- Profit-driven banks are less willing to mix $\rightarrow$ lower cutoff

Engagement on the bottom, Exit on the top

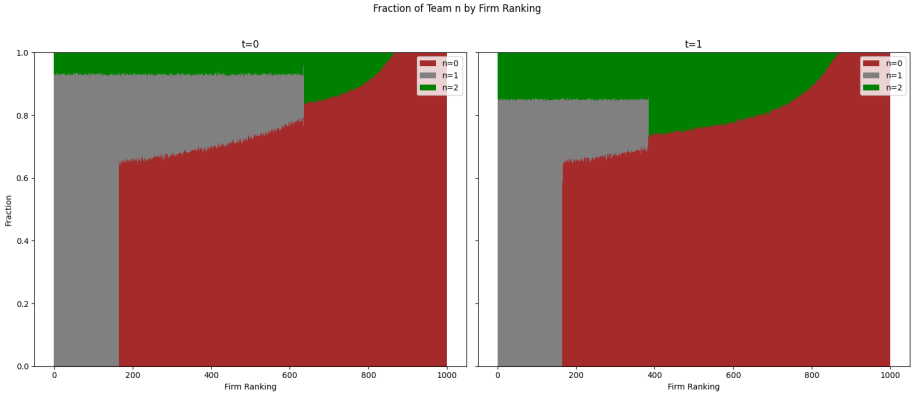


Figure 4: Firm matching outcomes before (left panel) vs. after the shock (right panel).

The Core Result: Exit vs. Engagement

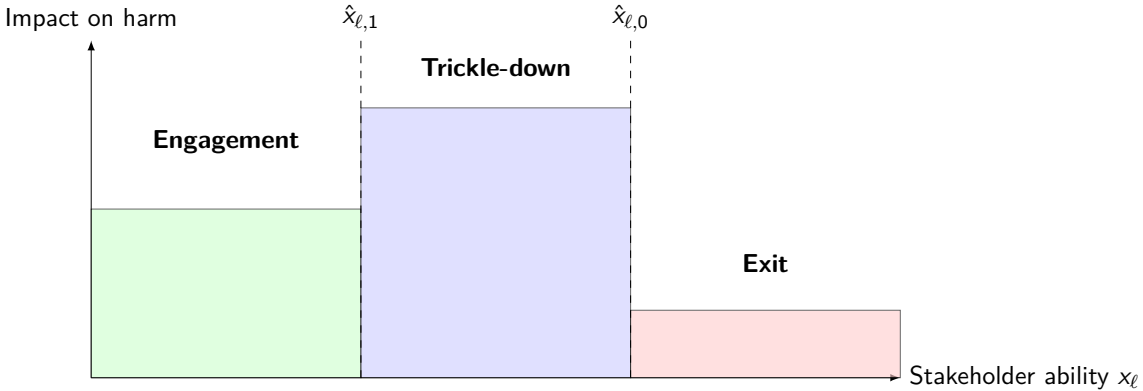
Proposition: Consider an increase in λ_ℓ (more purpose-driven type ℓ).

There exists a **cutoff firm size**:

- **Below cutoff:** Treated stakeholders **engage**
  - Stay with same-sized firm
  - Share mitigation costs
  - Compensating differentials adjust
- **Above cutoff:** Treated stakeholders **exit**
  - High-productivity firms outbid for scarce profit-driven stakeholders
  - Purpose-driven stakeholders reallocate to smaller, purpose-aligned firms
  - Exit appears ineffective at the firm level (moving to already-clean firms)

$\Rightarrow$ **Both exit and engagement** occur in equilibrium (vs. engagement-only under $\sigma = 0$)

Schematic: Micro-Level Impact of Shock



Impact zones of preference shock: engagement (bottom), trickle-down (middle), exit (top).

Impact Trickles Down

Proposition (Trickle-Down): The aggregate impact of exit has three regions:

1. **Top** ($x_\ell \geq \hat{x}_{\ell,0}$):
   - Some treated stakeholders have no direct effect
   - They exit to already-clean firms
   - But their displacement creates ripple effects in the middle

2. **Middle** ($\hat{x}_{\ell,0} \leq x_\ell \leq \hat{x}_{\ell,1}$):
   - Harm reduction $>$ treated share
   - **Displacement effect:** untreated purpose-driven stakeholders got pushed down

3. **Bottom** ($x_\ell \leq \hat{x}_{\ell,1}$):
   - Harm reduction $=$ treated stakeholder share (direct engagement)

The Trickle-Down Mechanism

Chain of displacement:

1. Large purpose-driven bank **exits** dirty firm $\rightarrow$ joins clean firm
2. Displaces smaller green bank at clean firm
3. Displaced bank moves to mid-tier firm
4. Mid-tier firm now has a purpose-driven bank $\rightarrow$ starts mitigating
5. This firm was **not directly connected** to the treated banks

$\Rightarrow$ Exit appears ineffective at the directly affected firm
$\Rightarrow$ But creates mitigation at firms **at lower ranking**

Firm-level studies miss this entirely

Implications for Empirical Research

- Standard empirical approach: **firm-level diff-in-diff**
  - Compare firms connected to treated vs. untreated banks
  - Finding: null or near-zero effect

- **Why this misses the aggregate impact:**
  - Treated firms at the top: brown banks substitute $\rightarrow$ no change
  - Impact occurs at **untreated mid-tier firms** through displacement
  - These firms are in the **control group** of the diff-in-diff!

- **Our model predicts:**
  - Treatment effect at treated firms ≈ 0 (consistent with data)
  - Aggregate effect $\gg 0$ (missed by firm-level studies)
  - Need **general equilibrium** analysis to capture full impact

GE Effect on Payoffs

- No changes in payoff iff the matching stays the same

$$U_\ell(x_\ell, \theta) = \int_{\underline{x}_\ell}^{x_\ell} z^*(\tilde{x}_\ell, \theta)\, d\tilde{x}_\ell + U_\ell(\underline{x}_\ell, \theta)$$

  - Compensating differential adjusts for stakeholers at the bottom
  - More mitigation $\Rightarrow$ **lower** compensation

- Who benefits from an increase in λ_ℓ on the top (exit)?

  - Remaining profit-driven banks (same side, different preference)
  - Purpose-driven workers at the top (different side, same preference)

Spillovers Across Stakeholder Groups

Proposition (Cross-Group Spillovers):

An increase in λ_ℓ (more purpose-driven type ℓ):

- **Lowers payoffs** for high-ability purpose-driven type-ℓ stakeholders
  - They are pushed to less productive teams
- **Raises payoffs** for high-ability purpose-driven other types $\ell' \neq \ell$
  - They now match with more productive purpose-driven type-ℓ stakeholders
- Example: More purpose-driven banks $\rightarrow$ **purpose-driven workers benefit**
  - They can now join teams with higher-ability green banks

Absent under $\sigma = 0$ — cross-group spillovers are a novel feature of our model

Application: Paris Agreement and Bank Decarbonization

Context: Post-2015 Paris Agreement

- Major banks commit to reduce fossil fuel exposure
- λ_b: 7% $\rightarrow$ 15% (share of purpose-driven banks)

Empirically,

1. Large banks **exit** (divest) rather than engage
2. Firm-level impact appears **limited** (null diff-in-diff)
3. Regional brown banks **substitute** for divesting green banks

Our model rationalizes all three:

1. Exit is optimal for large purpose-driven banks ($\sigma > 0$)
2. Firm-level null because impact operates through trickle-down
3. Brown bank substitution is the equilibrium response of dirty firms

Conclusion

1. **Exit vs. engagement** depends on whether harm scales with productivity
2. When $\sigma > 0$:
   - Large stakeholders exit $\rightarrow$ sort to purpose-aligned firms
   - Displacement creates **trickle-down** effects
   - Aggregate impact $\gg$ firm-level impact
3. **Cross-group spillovers:** Policies targeting banks also affect workers and suppliers
4. **Policy implication:** Don't dismiss exit as ineffective based on firm-level evidence
5. **Empirical implication:** Need GE approaches to measure true impact of exit

References I

Broccardo, E., Hart, O., Zingales, L., et al. (2022). Exit versus voice. *Journal of Political Economy*, 130(12):3101–3145.

Chiappori, P.-A., McCann, R., and Pass, B. (2016). Multidimensional matching. *arXiv preprint arXiv:1604.05771*.

Chiappori, P.-A., Oreffice, S., and Quintana-Domeque, C. (2018). Bidimensional matching with heterogeneous preferences: education and smoking in the marriage market. *Journal of the European Economic Association*, 16(1):161–198.

Dupuy, A. and Galichon, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, 122(6):1271–1319.

Gabaix, X. and Landier, A. (2008). Why has ceo pay increased so much? *The Quarterly Journal of Economics*, 123(1):49–100.

Hatfield, J. W. and Kominers, S. D. (2015). Multilateral matching. *Journal of Economic Theory*, 156:175–206. Computer Science and Economic Theory.

Heinkel, R., Kraus, A., and Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of financial and quantitative analysis*, 36(4):431–449.

Hong, H., Wang, N., and Yang, J. (2021). Welfare consequences of sustainable finance. Technical report, National Bureau of Economic Research.

Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies*, 84(2):718–789.

Oehmke, M. and Opp, M. M. (2023). A theory of socially responsible investment. *Swedish House of Finance Research Paper*, (20-2).

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571.

Pedersen, L. H., Fitzgibbons, S., and Pomorski, L. (2021). Responsible investing: The esg-efficient frontier. *Journal of Financial Economics*, 142(2):572–597.

Sattinger, M. (1979). Differential rents and the distribution of earnings. *Oxford Economic Papers*, 31(1):60–71.

Tervio, M. (2008). The difference that ceos make: An assignment model approach. *American Economic Review*, 98(3):642–68.