Stock Market Wealth and Entrepreneurship*

Gabriel Chodorow-Reich† Plamen T. Nenov‡ Vitor Santos§ Alp Simsek¶

May 2024

Abstract

We use data on stock portfolios of Norwegian households to show that stock market wealth increases entrepreneurship by relaxing financial constraints. Our research design isolates idiosyncratic variation in household-level stock market returns. An increase in stock market wealth increases the propensity to start a firm, with the response concentrated in households with moderate levels of financial wealth, for whom a 20 percent increase in stock wealth increases the likelihood to start a firm by about 20%, and in years when the aggregate stock market return in Norway is high. We develop a method to study the effect of wealth on firm outcomes that corrects for the bias introduced by selection into entrepreneurship. Higher wealth causally increases firm profitability, an indication that it relaxes would-be entrepreneurs' financial constraints. Consistent with this interpretation, the pass-through from stock wealth into equity in the new firm is one-for-one.

JEL Classification: E22, E44, L26, G50

Keywords: entrepreneurship, financial frictions, selection correction

*We would like to thank Søren Leth Petersen, Espen Moen, Lars Persson, Alex Popov, Kasper Roszbach, Martin Schmalz, Joacim Tåg, as well as seminar and conference participants at BI Norwegian Business School, Norges Bank, the Norges Bank Spring Institute, University of British Columbia, IFN Stockholm, University of Pittsburgh, Tufts University, the 2023 Leuven Summer Event, the 2023 Nordic Micro Data workshop, UC Berkeley, UT Austin, and ESSEC Paris, for valuable comments and suggestions. We also want to thank Janis Berzins and Øyvind Norli for providing us with the Oslo Børs data. This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank.

†Harvard University and NBER. Email: chodorowreich@fas.harvard.edu

‡Norges Bank and Norwegian Business School (BI). Email: plamen.nenov@norges-bank.no

§Norwegian Business School (BI). Email: vitor.santos@bi.no

¶Yale School of Management, NBER, and CEPR. Email: alp.simsek@yale.edu

1 Introduction

New business creation constitutes an integral part of economic growth and provides an important pathway to individual wealth accumulation (Buera et al., 2015; Smith et al., 2019; Bhandari et al., 2022). Does initial wealth affect who starts a firm or the type of business they build? In theory, higher wealth might increase business entry and profits by allowing would-be entrepreneurs to overcome fixed costs of set up. Yet, establishing a causal effect of wealth on entry faces the challenge that high wealth individuals may have other traits that make them more likely to start businesses (Hurst and Lusardi, 2004). Even granting a causal effect, determining why wealth matters or the effect of wealth on business outcomes such as profitability faces the further difficulty that an entry effect causes non-random selection on talent across entrepreneurs with different initial wealth (Buera, 2009).

We make four contributions. First, we study the effect of *stock market wealth*, the most volatile component of household financial wealth but one that has received comparatively little attention in the literature on entrepreneurship. Second, we show a causal effect of wealth on entry using a novel research design that compares individuals with the same *ex ante* stock portfolio characteristics but different *ex post* market returns arising from idiosyncratic holdings. Third, we isolate the causal effect of wealth on business outcomes by demonstrating a rank preservation property in a simple but general model of entrepreneurial choice. Fourth, in our data higher initial wealth causally increases key business variables, including firm profits and owners' equity, providing evidence that wealth matters to entrepreneurship by relaxing financial constraints.

To frame the empirical analysis, we start by describing the model and testable implications. Individuals differ in initial wealth and potential business productivity. Since our empirical approach isolates random variation in wealth, we assume that initial wealth is conditionally independent of productivity. An individual starts a business only if her utility as an entrepreneur exceeds her utility from wage employment. Greater wealth can affect this trade-off either by allowing for higher business profits (e.g. because of borrowing constraints) or by increasing the utility from running a business (e.g. because entrepreneurship provides non-pecuniary benefits). Comparing individuals with high and low initial wealth correctly identifies the magnitude of this entry effect (see Proposition 1).

The effect of wealth on business outcomes is harder to identify because we observe firm outcomes *only for entrants*, and entrants with different initial wealth have non-random differences in productivity. While this insight follows directly from a standard model of entrepreneurial selection, simple comparisons of outcomes of high and low wealth entrepreneurs remain prevalent in empirical work. We show that, under a natural monotonicity assumption, the model satisfies a rank preservation property that enables us to match entrants by productivity and estimate the causal effect of wealth on firm-level outcomes (see Proposition 2). Intuitively, we address non-random selection into entry by removing from the comparison the left tail of the productivity distribution of the firms with high initial wealth, as these firms would not have existed if the founder instead had low initial wealth. Propensity score reweighting extends this result to the realistic case of independence between wealth and ability *conditional on observables* (see Proposition 3). Finally, the model shows that the causal effect of initial wealth on firm profits provides a key moment to distinguish financial constraints from a non-pecuniary benefits explanation of the wealth effect on entry, as only under the former does higher wealth relax operating constraints and result in higher profits.

We bring the model's insights to the data using administrative records from Norway for both households and firms. We merge several administrative data sets, including a registry of security-level holdings of Norwegian stocks, total household financial wealth from tax records, labor market history from the employer-employee register, and firm balance sheet and income statements. In short, we observe household financial wealth, portfolio allocation at the individual stock level, business ownership, and firm-level outcomes. We define an entrepreneur as an individual who owns at least 1/3 of the book value of stocks in an incorporated non-financial firm, and where the firm has at most 3 stock owners. To remove shell companies we further require that the firm has at least one employee in the year of foundation or subsequent year and either holds no public equities or employs a worker who is not a member of the entrepreneur's household.

Our first main result compares entry into entrepreneurship across individuals with different stock market returns. There are two important threats to causal identification. First, realized stock market returns may correlate with other factors that affect the entrepreneurship decision. For example, home bias in portfolio choice could result in better returns in periods when the individual's industry or local area is booming. We solve this concern by including sector$\times$time and geography$\times$time fixed effects. Second, individuals more likely to become entrepreneurs might hold systematically different portfolios, for example if risk-tolerant individuals both hold riskier portfolios and are more likely to start firms. Following the results in Borusyak and Hull (2021), we address this concern by including flexible controls for *ex ante* portfolio characteristics such as the market beta. Our research design therefore isolates variation in market returns that comes from random realizations of the idiosyncratic component of portfolio holdings across portfolios with the same *ex ante* characteristics. This approach to isolating variation in returns by controlling for characteristics appears to be new to the literature analyzing the effects of stock market wealth (see e.g. Di Maggio et al., 2020; Ring, 2022). In our setting, more than half of households with directly held stocks have only one or two holdings, generating substantial idiosyncratic return variation.

We find strong evidence that greater stock market wealth increases entrepreneurship. Pooling across all observations, a 20 percentage point increase in total financial wealth due to an idiosyncratic stock return raises entry into entrepreneurship by about 1/10th of the sample average rate. This overall effect masks important underlying heterogeneity: individuals with sufficiently high initial wealth exhibit essentially no effect of higher wealth on entrepreneurship, while for lower wealth individuals the marginal effect is essentially doubled, with a 20 percentage point increase in wealth raising the entry rate by about 20%. We also find that the positive marginal effect comes only from years when the overall stock market does well. We demonstrate robustness along several dimensions, including controlling for initial wealth and labor income or for the sector of the individual's largest stock holding. The positive effect of wealth on entry also holds when restricting the sample to similarly under-diversified households, those that hold less than 3 stocks and whose main stock holding is among the 20 most popular companies among small investors traded on the Norwegian stock exchange. Placebo exercises demonstrate no response of entrepreneurship to future stock returns.

Our second set of results apply the model's selection correction to obtain the causal effect of wealth on firm outcomes. We find sizable positive effects of higher portfolio returns on firm balance sheet and income statement outcomes, including capital, sales, employment, value added, and profitability. As shown in our model, the increase in profits indicates a role for financial constraints. Two additional results further this interpretation: (i) a marginal increase in stock wealth results in a nearly one-to-one increase in owners' equity in the new firm, and (ii) entrepreneurs with higher returns finance their larger businesses by actively liquidating stocks.

Together, our results provide evidence that stock market wealth increases entrepreneurship and results in more profitable firms at creation. This evidence complements research using other sources of windfall gains such as housing capital gains or lottery winnings. Understanding the effects of stock market fluctuations in particular are important because the stock market accounts for a large share of total wealth fluctuations. Furthermore, stock market wealth is special in the sense that it rises when the stock market does well, which our evidence shows is also exactly when such wealth matters most to entrepreneurship. Finally, the relaxation of financial constraints provides a key transmission channel for these results.

Related literature. Our paper makes a direct contribution to the literature on wealth and entrepreneurship and in particular to the academic debate on the importance of liquidity constraints for business creation.[1] In two early seminal contributions, Evans and Jovanovic (1989) and Evans and Leighton (1989) find a positive association between individual wealth and the propensity to start a business using data from the National Longitudinal Survey of Youth. Using a structural approach, Evans and Jovanovic (1989) argue that this relationship is likely causal. Similar correlational findings have been reported in Blanchflower and Oswald (1998), Fairlie (1999), Quadrini (1999), and Gentry and Hubbard (2004) among others.

On the other hand, Hurst and Lusardi (2004) find a flat relationship in the Panel Study of Income Dynamics between wealth and entry into entrepreneurship for most of the wealth distribution and a strong positive relationship only at the top. In addition, they find no evidence that individual wealth matters more for entry in high starting-capital industries. More recently, Bhandari et al. (2022) use administrative data from the IRS and Social Security Administration and find that entrants into self-employment have *lower* asset incomes prior to entry.

The lack of consensus on the relationship between wealth and entrepreneurship spurred a literature that looks for exogenous shocks to wealth. Several studies have found that individuals receiving an inheritance are more likely to become entrepreneurs (Holtz-Eakin et al. (1994), Lindh and Ohlsson (1996), Blanchflower and Oswald (1998), Andersen and Nielsen (2012), Fairlie and Krashinsky (2012)). Hurst and Lusardi (2004) challenge this approach by showing that both past and *future* inheritances predict entry into entrepreneurship, suggesting that inheritance may correlate with other factors such as risk tolerance or preferences that determine entry.

Another form of variation in wealth comes from lottery winnings or other cash windfalls. Lindh and Ohlsson (1996) report a positive effect of lottery winnings on firm creation in Sweden while Cesarini et al. (2017) find that lottery winnings reduce self-employment income in Swedish data. Cespedes et al. (2021) investigate the effect of a retail business receiving a bonus payment for selling a winning jackpot ticket and find both an intensive and extensive (serial entrepreneurship) effect that depends on the size of the lottery windfall. Using U.S. administrative tax records, Golosov et al. (2021) find a positive effect of lottery winnings on the propensity to report small ($< \$15,000$) self-employment income but no effect on transitioning to more substantial business activity. Bermejo et al. (2022) find a positive effect of winning the Spanish Christmas lottery on regional firm creation, which they argue

[1]Parker (2018) provides a comprehensive treatment of the economics of entrepreneurship including the link between wealth and entrepreneurship. See also Kerr and Nanda (2011). In addition to empirical studies, the idea of financial constraints impacting entry into entrepreneurship has also been explored theoretically and quantitatively in Banerjee and Newman (1993), Aghion and Bolton (1997), Piketty (1997), Cagetti and De Nardi (2006), Buera and Shin (2013), and Moll (2014), among others. See Buera et al. (2015) for a review of that literature.

is mediated via a credit constraint channel. Finally, Bellon et al. (2021) find a positive effect on incorporated business creation from cash windfalls due to shale oil exploration contracts in Texas.

The third wealth shock used in the literature is to housing wealth (Hurst and Lusardi (2004), Adelino et al. (2015), Corradin and Popov (2015), Schmalz et al. (2017), Jensen et al. (2022), Kerr et al. (2022)). With the prominent exception of Hurst and Lusardi (2004), most of this literature finds that housing wealth increases business creation. However, since variation in housing wealth is mostly regional there are inherent difficulties in establishing whether these effects are driven by local economic shocks, by higher wealth, or by higher collateral values. Schmalz et al. (2017) cleverly circumvent this difficulty by comparing local homeowners to renters and find that local homeowners are more likely to start a business compared to renters after a local house price appreciation. They also compare local homeowners with and without a mortgage on their house and show that the effects are present only for homeowners without a mortgage – a sign of higher housing wealth relaxing liquidity constraints, since in France homeowners with a mortgage cannot take on more mortgage debt.[2]

Relative to this large existing literature, and to the best of our knowledge, we are the first to propose and implement an empirical design featuring the impact of stock market wealth on business creation.[3] Furthermore, our administrative data contain the near-universe of households and firms over a long time period with no top-coding, non-observability of assets, or self-reporting errors that have been an issue for much of the existing literature. Our main findings point to a robust causal and economically significant effect of stock market wealth on business creation.

Several earlier papers also examine the effect of wealth on firm-level outcomes. Using data on start-ups from Norway, Hvide and Møen (2010) find a negative relationship between wealth and start-up profitability at the top of the wealth distribution. Similarly, Andersen and Nielsen (2012) find that on average the firms created with unexpected inheritances have lower survival and profitability. Much of the rest of the literature reaches the opposite conclusion. Holtz-Eakin et al. (1994) show that firms whose owners receive inheritances tend to survive longer. Schmalz et al. (2017) find that firms created by homeowners in periods of rising house prices are significantly larger at the time of creation than those started by renters and that such firms tend to survive longer and are therefore not "riskier" in the sense of having a higher probability of failure. Jensen et al. (2022) also find higher survival

[2]In addition to collateral shocks some papers have considered the effects of credit market shocks on entrepreneurship (Black and Strahan (2002), Kerr and Nanda (2009), Fracassi et al. (2013), among others).

[3]In contemporaneous and complementary work, Chetty et al. (in progress) show that early employees at firms undergoing an IPO have higher subsequent entrepreneurship rates. Our work differs in focusing on more "ordinary" stock market participants who have not necessarily already joined newly-formed firms.

rates among firms created with higher housing wealth. McKenzie (2017) uses the random allocation of grants to business start-ups in Nigeria to show that both potential and existing entrepreneurs that receive a grant are both more likely to operate a business three years after the grant allocation and more likely to operate firms with ten or more workers.[4] Bermejo et al. (2022) find a positive and significant effect on firm size and survival using the Spanish Christmas lottery.

None of these papers attempts to separately identify the causal effect of wealth conditional on becoming an entrepreneur from the changing distribution of entrants induced by higher wealth.[5] Indeed, our selection model can help to reconcile the disparate findings across papers, since it implies that unconditional differences between firms started by high and low wealth entrepreneurs have theoretically ambiguous sign and could vary across institutional settings. Our selection correction echoes the structural approach in Evans and Jovanovic (1989), but provides a more direct mapping from data to results. After applying our selection correction, we find strong evidence that higher wealth causes better firm outcomes.

The literature has also debated whether to interpret a causal relationship between wealth changes and entrepreneurship as revealing liquidity constraints. For example, higher wealth could make individuals more risk tolerant or overly optimistic and, hence, also more likely to accept the non-diversification risk of starting a business (Kihlstrom and Laffont, 1979; Moskowitz and Vissing-Jørgensen, 2002; Landier and Thesmar, 2008; Hall and Woodward, 2010). Or there might be a *non-pecuniary benefit* from firm ownership that increases with wealth (Hurst and Pugsley, 2017).[6] In this spirit, Hamilton (2000) finds a "self-employment" discount, showing that entrepreneurship tends to persist despite lower earnings growth compared to paid employment (see also Pugsley and Hurst (2011) and Catherine (2022)), although Bhandari et al. (2022) challenge this result using their administrative data.

Our model demonstrates that the causal effect of wealth on firm profits can help to distinguish among these channels. Our finding of higher profits suggest wealth increases entrepreneurship at least in part by relaxing financial constraints. Consistent with this interpretation, we also find a near one-to-one pass-through from wealth to equity in the firm and that stock market wealth matters less for entrepreneurship decisions of high financial wealth households. A finding that financial constraints matter in turn creates a possible role

[4]See also De Mel et al. (2008) who show that wealth shocks are important for the performance of microenterprises in Sri Lanka.

[5]Buera (2009) notes that selection implies that the distribution of entrepreneurial ability among workers and entrepreneurs varies along the wealth distribution, but in the context of the relationship between wealth and entrepreneurship rather than the effect of wealth on business outcomes.

[6]An alternative explanation for entrepreneurs investing substantial equity and thus holding substantial non-diversified risk in their firms focuses on signaling incentives to lenders under asymmetric information, see Leland and Pyle (1977), Ross (1977), and Nenov (2017), among others.

for development policies such as business subsidies that mitigate financial frictions (Itskhoki and Moll, 2019).

Our focus on the effect of stock market wealth for the real economy brings our paper close to recent work by Di Maggio et al. (2020), Chodorow-Reich et al. (2021) and Ring (2022), among others. Unlike these papers, we consider the effects of stock market wealth for entrepreneurship and business creation rather than the effects on consumer spending (Di Maggio et al., 2020; Chodorow-Reich et al., 2021) or the effects of wealth of existing firm owners on firm employment and investment during the 2008-2009 financial crisis (Ring (2022)). We also contribute to this literature by formalizing the problem of non-random portfolio formation raised by Borusyak and Hull (2021).

2 The model and testable hypotheses

We start by presenting a simple but general model of entry into entrepreneurship. We use the model to illustrate the hypotheses that we test in the empirical analysis. We also highlight how endogenous entry into entrepreneurship introduces selection on unobserved productivity and characterize a procedure to correct for this selection effect.

2.1 Setup

The model has a single period (we abstract from dynamics). There is a continuum of agents denoted by i with mass normalized to one. Agents are associated with unobservable potential business productivity $z_i \in Z \subset R_+$, observable initial assets $a_i \in A \subset R_+$, and an observable vector of covariates $x_i \in X$ that might be correlated with productivity and assets. Let $F_{za}(z_i, a_i|x_i)$ denote the joint cumulative distribution function (CDF) of productivity and assets conditional on x_i. We assume the corresponding probability distribution function (PDF), denoted by $f_{za}(z_i, a_i|x_i)$, is continuous. We use a similar notation for other distributions. For instance, $F_z(z_i|a_i, x_i)$ denotes the CDF of z_i conditional on a_i and x_i, and $F_z(z_i|x_i)$ denotes the marginal CDF of z_i conditional on x_i.

In the empirical analysis, we assume $a_i = a_{i,t-1}r_{i,t}$ and isolate quasi-random fluctuations in the portfolio return $r_{i,t}$ by controlling for covariates such as *ex ante* portfolio characteristics. Therefore, in the model we impose the following conditional independence assumption:

Assumption (CIA). z_i and a_i are independent *conditional on* x_i, that is: $F_z(z_i|a_i, x_i) = F_z(z_i|x_i)$ (and a similar condition holds for F_a).

An individual i chooses whether to enter into business, $E_i \in \{0, 1\}$. If she does not enter, $E_i = 0$, she earns the outside option (reservation wage) $w(z_i)$. The reservation wage can depend on z_i, but we require this dependence to be relatively small in a sense that we formalize below. In this case, the individual's (consumption-equivalent) utility is equal to her assets plus her wage

$$U\left(E_i = 0\right) = a_i + w\left(z_i\right).$$

If instead the individual enters, $E_i = 1$, she runs a business with size $k_i = k\left(z_i, a_i\right)$ and earns profits $\pi_i = \Pi(k_i; z_i, a_i) = \pi\left(z_i, a_i\right)$. The individual's consumption-equivalent utility is

$$U\left(E_i = 1\right) = a_i + \pi\left(z_i, a_i\right) + u^e\left(z_i, a_i\right).$$

Here, the profit term captures the pecuniary benefit from entrepreneurship and the last term captures a possible non-pecuniary benefit—individuals might enjoy running their own business. An individual enters into business if her potential consumption-equivalent utility from entrepreneurship exceeds her potential wage

$$U\left(E_i = 1\right) \geq U\left(E_i = 0\right) \Longrightarrow \pi\left(z_i, a_i\right) + u^e\left(z_i, a_i\right) \geq w(z_i).$$

Observe that we assume the outside option w, size k, profit π, and entrepreneurship benefit u^e depend *only* on productivity and initial wealth. Observed covariates x_i affect these outcomes only through their impact on z_i, a_i and other unobserved variables do not affect these outcomes. These assumptions are restrictive but they help to illustrate our results while simplifying the notation. In Appendix B.2, we show that under additional assumptions our results extend to cases in which the outcomes w, k, π, u^e can be heterogeneous in other (observed or unobserved) dimensions. We discuss this extension in Section 2.4.

To characterize and estimate the effects of wealth on entrepreneurship, we impose mild monotonicity conditions on the profit, entrepreneurship benefit, and size functions.

Assumption (M). $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{da_i} \geq 0$ and $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i)-w(z_i))}{dz_i} > 0$, $\frac{dk(z_i,a_i)}{dz_i} > 0$.

The first condition captures the main effect we investigate: greater wealth increases the total utility from entrepreneurship. This could be either because greater wealth increases profits by relaxing financial constraints, or because greater wealth increases the non-pecuniary benefits from entrepreneurship. We will shortly show how to distinguish among these explanations. The last two conditions enable us to address the selection effect induced by entry and estimate the effect of wealth on firm-level outcomes. The second condition says that productivity increases the total utility from entrepreneurship faster than it increases the reservation wage (so it increases the net gain from entrepreneurship). The third condition says that individuals with higher productivity start larger businesses.

2.2 Effect of Wealth on Entry

Our first goal is to understand how initial wealth affects the entrepreneur's propensity to enter into business. To this end, we define the fraction of agents with assets a that enter into business conditional on covariates

$$e\left(a|x_i\right) = \int_{(z_i,a)} E_i dF_{za}\left(z_i, a|x_i\right) \in (0,1)\,.$$

Our first result shows that this fraction is increasing in initial assets: greater wealth increases the propensity to enter. This is the first hypothesis that we test in our empirical analysis.

Proposition 1 (Causal effect of assets on entry)**.** *Consider the entry model with Assumption (M). Given a, there exists a threshold productivity level $\overline{z}\left(a\right)$ such that an agent enters iff $z \geq \overline{z}\left(a\right)$. The threshold productivity $\overline{z}\left(a\right)$ is* weakly decreasing *in initial assets. Higher assets induce higher entry into business: the fraction of entrants conditional on covariates, $e\left(a|x_i\right)$, is given by $\int_{z_i \geq \overline{z}(a)} dF_z\left(z_i|x_i\right)$, and it is weakly* increasing *in initial assets a.*

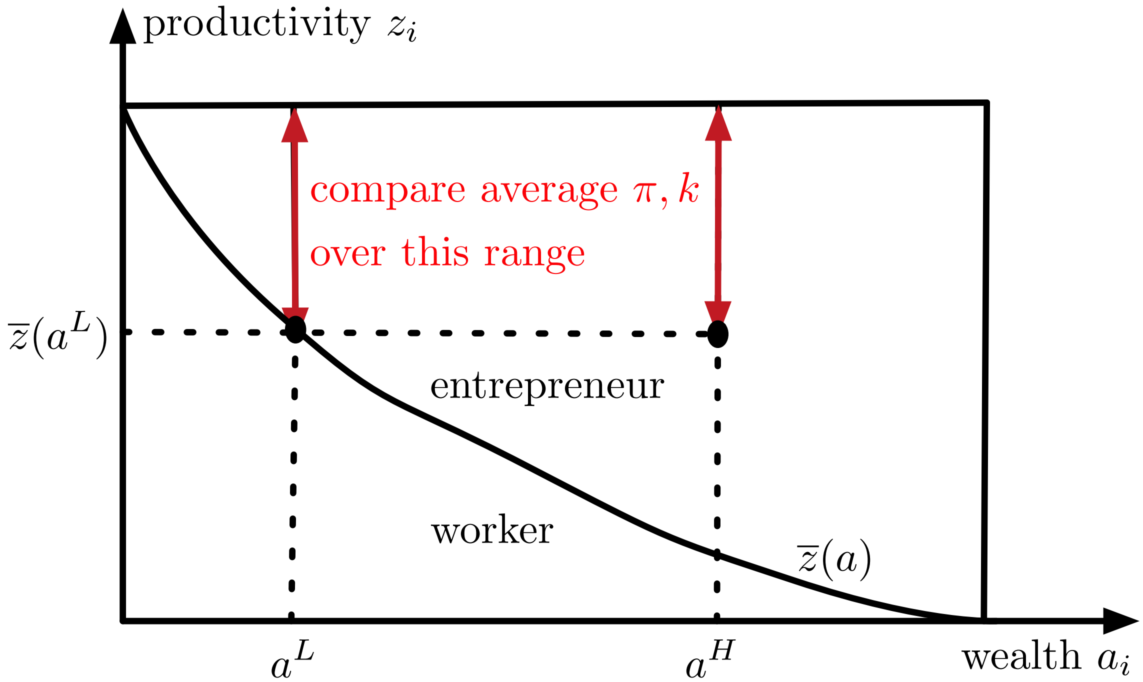
Under Assumption (M), individuals enter only if they have either sufficiently high wealth or sufficiently high productivity. Thus, there is a threshold productivity $\overline{z}\left(a\right)$, *decreasing in wealth*, such that individuals enter only if their productivity exceeds this level. This also implies that, conditional on covariates, greater wealth increases the propensity to enter. Figure 1 illustrates the threshold function $\overline{z}\left(a\right)$ and the region of entry in a particular example.

2.3 Effect of Wealth on Business Characteristics

Our second goal is to understand how initial wealth affects initial firm-level outcomes, such as firm profits, conditional on entry. This exercise is more complicated because entry into business induces selection on *unobserved* productivity. To see this, consider the *average* firm profits given wealth (and covariates) and *conditional on entry*,

$$\begin{aligned} \pi\left(a|x_i\right) &= \mathbb{E}\left[\pi_i|x_i, a_i = a, E_i = 1\right] \\ &= \frac{\int_{z_i \geq \overline{z}(a)} \pi\left(z_i, a\right) dF_z\left(z_i|x_i\right)}{e\left(a|x_i\right)} \text{ where } e\left(a|x_i\right) = \int_{z_i \geq \overline{z}(a)} dF_z\left(z_i|x_i\right). \end{aligned}$$

Figure 1: A Model of Endogenous Entry into Business



Notes: Individuals above the curved line, $\overline{z}(a)$, enter into business. We identify the causal effect of wealth on profits and capital by comparing the outcomes for individuals that would have entered both with low wealth a_L and with high wealth a_H.

Suppose we empirically estimate $\frac{d\pi(a|x_i)}{da}$. This does not necessarily identify $\frac{d\pi(z_i,a|x_i)}{da}$, because changing assets a also affects the average productivity of entrants. In our model, increasing a *reduces* the average productivity of entrants (see Figure 1).

Since we do not directly observe productivity, we cannot address this selection problem by controlling for productivity z. Instead, we observe the initial wealth a for all individuals and capital k (or firm size) for the individuals that enter. Our second result shows how to use these *observed* outcomes to match individuals by productivity and estimate firm-level outcomes.

The result relies on sorting the entrants by their size k—which is monotonically increasing with productivity in view of Assumption (M)—to match them according to their unobserved productivity. To formalize this idea, we define the fraction of individuals with wealth a that enter into business *and whose size exceeds a cutoff level* $\overline{k}$ conditional on covariates:

$$e\left(a, k \geq \overline{k}|x_i\right) = \int_{z_i \geq \overline{z}(a), k(z_i,a) \geq \overline{k}} dF_z\left(z_i|x_i\right).$$

In principle, we can observe these fractions conditional on x_i for any assets a and threshold $\overline{k}$. We also let $y\left(z_i, a_i\right)$ denote a firm-level outcome that is observed and that can be described as a function of the entrant's productivity and initial wealth, such as profits $\pi\left(z_i, a\right)$ or size $k\left(z_i, a\right)$.

Proposition 2 (Rank preservation and the causal effect of assets on firm-level outcome)**.** *Consider an entrant with covariates x_i, initial wealth a^L, size k^L, and firm-level outcome y^L, along with* unobserved *productivity $z \geq \overline{z}\left(a^L\right)$. Let $a^H > a^L$ denote a higher wealth level and $\overline{k} \geq k^L$ denote the unique solution to the following:*

$$e\left(a^H, k \geq \overline{k}|x_i\right) = e\left(a^L, k \geq k^L|x_i\right). \tag{1}$$

Let $\overline{y}$ denote the outcome corresponding to the firm with higher initial wealth a^H and the cutoff size $\overline{k}$. Then, $\overline{y} = y\left(z, a^H\right)$: that is, $\overline{y}$ is the firm-level outcome the entrant would have if she had higher initial wealth (and the same productivity). Thus, the difference

$$\overline{y} - y^L = y\left(z, a^H\right) - y\left(z, a^L\right)$$

identifies the causal effect of initial wealth on the firm-level outcome for an entrant with productivity z.

We refer to condition (1) as *rank preservation*. To understand this condition, consider an entrant with covariates x_i, wealth a^L, and unobserved productivity z, and suppose we increase her wealth from a^L to a^H. In view of assumption (M), this change would leave the entrant's *relative rank for size* unchanged. Intuitively, with either a^L and a^H, the individuals with productivity $\tilde{z} \geq z$ also enter and have a greater size than the entrant (and the individuals with $\tilde{z} < z$ would either not enter, or they would enter and have a smaller size than the entrant). Importantly, since we observe the size by rank for each asset level (captured by the functions $e\left(a, k \geq \overline{k}|x_i\right)$), we can solve (1) and calculate the size that the entrant would have if she had higher initial wealth. This in turn enables us to estimate the causal effect of initial wealth on size as well as other firm-level outcomes such as profits.

While we can in principle compute the fractions $e\left(a, k \geq \overline{k}|x_i\right)$ and implement Proposition 2 separately for each x_i, in practice this computation is not feasible because even large population data sets, such as the one we use, have a much smaller number of entrepreneurs. Therefore, we also develop a version of the proposition that uses *unconditional* fractions along with propensity score reweighting to control for the covariates x_i. To state the result, consider the fraction of entrepreneurs with wealth a and minimum size $\overline{k}$ aggregated over all covariates:

$$e\left(a, k \geq \overline{k}\right) = \int_{x_i} \int_{z_i \geq \overline{z}(a), k(z_i, a) \geq \overline{k}} dF_z\left(z_i | x_i\right) dF_x\left(x_i | a\right).$$

Matching $e\left(a^L, k^L\right)$ with $e\left(a^H, k \geq \overline{k}\right)$ (with appropriate $\overline{k}$) will no longer control for productivity since agents with different wealth levels can be associated with different covariates (captured by $dF_x\left(x_i|a\right)$) and these covariates can be associated with different levels of productivity (captured by $dF_z\left(z_i|x_i\right)$). However, we can correct for these differences by appropriately reweighting the marginal distributions $dF_x\left(x_i|a\right)$, following the large literature on propensity score reweighting (see, e.g., Rosenbaum and Rubin (1983); DiNardo et al. (1996); Heckman et al. (1998); Hirano et al. (2003)).

Formally, for the lower wealth level a^L, we define the *reweighted* fraction of entrepreneurs:

$$e^*\left(a^L, k \geq \overline{k}\right) = \int_{x_i} \int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq \overline{k}} dF_z\left(z_i | x_i\right) \omega\left(x_i\right) dF_x\left(x_i | a^L\right) \tag{2}$$

$$\text{where } \omega\left(x_i\right) = \frac{dF_x\left(x_i | a^H\right)}{dF_x\left(x_i | a^L\right)} = \frac{dF_a\left(a^H | x_i\right)}{dF_a\left(a^L | x_i\right)} \frac{dF_a\left(a^L\right)}{dF_a\left(a^H\right)}.$$

The second line defines the propensity weights $\omega\left(x_i\right)$ and applies Bayes rule. These weights can be estimated from data since they rely *only* on observable variables and are defined over the full set of agents. Intuitively, the fraction $e^*\left(a^L, k \geq \overline{k}\right)$ overweights (resp. underweights) the agents with covariates x_i that are relatively more common (resp. less common) among agents with higher wealth a^H. This reweighting makes the sample with a^L comparable to the sample with a^H in terms of the distribution of covariates. Consequently, a version of Proposition 2 applies with the reweighted distribution.

Proposition 3 (Rank preservation with propensity score reweighting)**.** *Consider the entrants with initial wealth a^L, size k^L, and firm-level outcome y^L, along with* unobserved *productivity $z \geq \overline{z}\left(a^L\right)$. Let $a^H > a^L$ denote a higher wealth level and $e^*\left(a^L, k \geq \overline{k}\right)$ denote the propensity score reweighted fraction of entrants with lower wealth a^L defined in* (2)*. Let $\overline{k} \geq k^L$ denote the unique solution to:*

$$e\left(a^H, k \geq \overline{k}\right) = e^*\left(a^L, k \geq k^L\right). \tag{3}$$

Let $\overline{y}$ denote the outcome corresponding to the firm with higher initial wealth a^H and the cutoff size $\overline{k}$. Then, $\overline{y} = y\left(z, a^H\right)$: that is, $\overline{y}$ is the firm-level outcome the entrant would *have if she had higher initial wealth (and the same productivity). Thus, the difference*

$$\overline{y} - y^L = y\left(z, a^H\right) - y\left(z, a^L\right)$$

identifies the causal effect of initial wealth on the firm-level outcome for an entrant with productivity z.

This result shows how to estimate the causal effect for particular entrants (with a specific size and productivity). In the empirical analysis, we focus on the *average* causal effect over all entrants with some initial wealth. Formally, fix a wealth level a^L and let k^L denote the *lowest-size* firm corresponding to entrants with a^L, with productivity given by the entry threshold, $z = \overline{z}\left(a^L\right)$. Eq. (3) then becomes:

$$e\left(a^H, k \geq \overline{k}\right) = e^*\left(a^L\right) = \int_{x_i}\int_{z_i \geq \overline{z}(a^L)} dF_z\left(z_i|x_i\right)\omega\left(x_i\right)dF_x\left(x_i|a^L\right).$$

In particular, among the entrants with higher level of assets a^H, we find the lowest-size entrant that would have entered also with the lower level of assets a^L, after balancing the covariates with propensity score reweighting. Denote the cutoff size with $\overline{k} = k^H\left(a^L\right)$. We then calculate the average outcome variable for high-asset entrants *with sizes above the cutoff*

$$\mathbf{y}\left(a^H, k \geq k^H\left(a^L\right)\right) = \mathbb{E}\left[y_i|a_i = a^H, E_i = 1, k \geq k^H\left(a^L\right)\right]. \tag{4}$$

We also calculate the average outcome variable for low-asset entrants *after propensity-score reweighting*

$$\mathbf{y}^*\left(a^L\right) = \mathbb{E}^*\left[y_i|a_i = a^L, E_i = 1\right] \equiv \frac{\int_{x_i}\int_{z_i \geq \overline{z}(a^L)} y\left(z_i, a^L\right)dF_z\left(z_i|x_i\right)\omega\left(x_i\right)dF_x\left(x_i|a^L\right)}{e^*\left(a^L\right)}. \tag{5}$$

Using Proposition 3, it is then easy to check that comparing $\mathbf{y}\left(a^H, k \geq k^H\left(a^L\right)\right)$ and $\mathbf{y}^*\left(a^L\right)$ identifies the average causal effect among the agents with high wealth a_H and productivity $z \geq \overline{z}\left(a^L\right)$.[7]

Figure 1 illustrates this approach. By considering the high-wealth entrants with size above the cutoff, we select individuals with relatively high productivity $z_i \geq \overline{z}\left(a^L\right)$. These

[7]In particular, we have

$$\mathbf{y}\left(a^H, k \geq k^H\left(a^L\right)\right) - \mathbf{y}^*\left(a^L\right) = \frac{\int_{x_i}\int_{z_i \geq \overline{z}(a^L)}\left(y\left(z_i, a^H\right) - y\left(z_i, a^L\right)\right)dF_z\left(z_i|x_i\right)dF_x\left(x_i|a^H\right)}{e^*\left(a^L\right)}.$$

individuals enter regardless of whether they start with wealth a^L or a^H: they are not subject to the selection effect we mentioned earlier. Therefore, comparing their average outcomes identifies the average causal effect of wealth on firm-level outcomes. In our empirical analysis, we calculate the empirical counterparts to $(4-5)$ and report the difference.

2.4 Residual Heterogeneity

Underlying the rank preservation condition in Propositions 2 and 3 is an inversion of the ranking of observed business size k to infer the ranking of unobserved productivity z. This inversion explains why we assume that k depends only on z and wealth a. In practice, k can also be heterogeneous along other dimensions. For instance, *ex post* heterogeneity in size (or profits) could arise because an individual might be subject to productivity shocks after deciding to enter into business or because she might make a mistake (relative to the optimal choice). In addition, *ex ante* heterogeneity in size (or profits) can emerge as individuals might have industry-specific skills that imply variation in their prospective firms' production processes or startup costs, adding an additional argument to the functions determining k (and π). Likewise, individuals' outside options w or their utility from entrepreneurship u^e might feature residual heterogeneity that is not fully captured by productivity z, e.g., due to differences in potential wage income and entrepreneurial productivity (or taste). These observations raise the question of how these types of unobserved residual heterogeneity affect our results.

In Appendix B.2, we show that our rank preservation approach is robust to allowing for ex-post and ex-ante residual heterogeneity under two additional assumptions. First, we require the residual heterogeneity to be independent from initial wealth and entrepreneurial productivity conditional on the observed covariates x. Second, we focus on entrants with size levels k that exceed the entry cutoff for size by a sufficient margin.

The intuition for these conditions is as follows. While residual heterogeneity shuffles the firms' sizes, it does not necessarily bias our approach in a particular direction when it is conditionally independent from initial wealth and productivity. Some less productive firms become larger and some more productive firms become smaller, but on average the firms with a given size k have a similar productivity as in the baseline without residual heterogeneity. In fact, under appropriate technical assumptions, the model satisfies certainty-equivalence properties that enable us to extend our main result to the case with residual heterogeneity as long as we focus on individuals that are sufficiently far from the entry cutoffs (see Proposition 4).

For individuals close to the entry cutoffs, our approach is not necessarily robust to residual heterogeneity, because of a selection problem induced by the interaction of unobserved residual heterogeneity with the entry decision. Some individuals that would choose size k have relatively low productivity (and high residual-induced size) so they might choose not to enter. Therefore, *the entrants* with size k might have a higher average productivity than in the baseline case. Since this selection might be different for the control group of individuals with lower initial wealth and the test group with higher initial wealth, our rank preservation approach does not necessarily control for the average productivity between these two groups. We alleviate this concern in two ways. First, we present robustness exercises where we focus on entrant firms whose size exceeds the entry cutoff by some margin. For a sufficiently large margin (that depends on the extent of unobserved residual heterogeneity), these firms are not subject to the selection concern driven by entry (see Appendix B.2). Second, for our baseline analysis without a margin, we note that this issue is likely to bias our results such that the test group with higher initial wealth will have a lower average productivity than the control group with lower initial wealth. This is because high initial wealth increases the propensity to enter (as we will verify empirically) and mitigates the selection concerns driven by the interaction of unobserved residual heterogeneity and entry. Since productivity tends to increase firm size and profits, the actual effects of wealth on these types of outcomes are likely to be larger than our estimates, that is, our estimates for firm-level outcomes can be viewed as a lower bound.

We further alleaviate the concerns with residual heterogeneity with additional robustness exercises that we describe in Section 6.3. Finally, it merits emphasizing that Proposition 1 does not depend on the inversion property. As a result, it goes through with any form of heterogeneity as long as Assumption (CIA) holds.

2.5 Explicit Microfoundations and the Role of Profits

In Appendix B.1 we describe two example models that satisfy Assumption (M) and hence in which Propositions 1 to 3 hold. These models illustrate complementary mechanisms by which wealth might increase entrepreneurship, while making distinct predictions for how wealth affects profits, $\frac{d\pi(z_i,a_i)}{da_i}$. The first model features financial frictions but does not have non-pecuniary benefits from entrepreneurship, $u^e(z_i, a_i) = 0$. In that model, the entrepreneur needs to obtain financing to pay for the fixed cost of starting a business and for the capital expenditures. She can borrow funds from outside financiers, but outside financing is costly and these costs are increasing in the amount that the entrepreneur borrows. Higher initial wealth (internal funds) reduces the need for outside financing, which in turn raises the entrepreneur's potential profits, $\frac{d\pi(z_i,a_i)}{da_i} \geq 0$, firm scale, $\frac{dk(z_i,a_i)}{da_i} \geq 0$, and the total utility from entrepreneurship, $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{da_i} \geq 0$.

The second model features no financial frictions but the non-pecuniary benefits from entrepreneurship are given by a function $u^e = U^e(k, c; z_i, a_i)$ where k is the size of the business, $c = a_i + \Pi(k; z_i)$ is regular consumption, and $\Pi(k; z_i)$ denotes profits as a function of size. We assume $\frac{dU^e}{dc} > 0$, $\frac{dU^e}{dk} \geq 0$, $\frac{d^2U^e}{dcdk} \geq 0$ (along with standard regularity conditions). These assumptions capture the idea that individuals enjoy running a (larger) business, and more so when their regular consumption is higher. Individuals choose their business size k to maximize $\Pi + U^e$. In this case, greater wealth increases the total utility from entrepreneurship as before, $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{da_i} \geq 0$, but the effect works through the non-pecuniary benefits $u^e(z_i, a_i)$. Crucially, unlike in the financial frictions model greater wealth *decreases* profits, $\frac{d\pi(z_i,a_i)}{da_i} \leq 0$. Intuitively, the non-pecuniary benefits from entrepreneurship induce a firm size beyond the profit-maximizing level, which reduces profits. Our empirical implementation of Propositions 2 and 3 allows us to estimate the effect of wealth on profits $\frac{d\pi(z_i,a_i)}{da_i}$, which we will use (along with other findings) to differentiate between the two models.[8]

3 Data and Definitions

We combine a number of administrative data sets from Norway using unique personal and firm identification numbers as well as the unique ISINs of publicly traded shares. The unit of analysis is a household. Our data set construction shares similarities with Fagereng et al. (2020) and Ring (2022).

3.1 Data

We obtain information on the composition of stock portfolios and business ownership from the shareholder register ("Aksjonærregisteret"). This data set records information on ownership of shares in Norwegian limited liability firms, both publicly traded and privately held, and the book value of those shares at the end of each calendar year starting in 2004. The information is collected by the Norwegian tax authority and is third-party reported by financial intermediaries and includes stocks held in individual retirement accounts.[9] Using the

[8]Both models satisfy the remaining two conditions in Assumption (M) when the reservation wage are constant (or only weakly increasing in productivity). This is because the total utility from entrepreneurship and the size of the business are both increasing with productivity—these are standard effects that continue to hold with financial frictions or with non-pecuniary utility from entrepreneurship.

[9]Most directly-held stock wealth in Norway is held outside of individual retirement accounts, as such accounts were not particularly wide-spread for most of our sample. Smogeli and Halvorsen (2019) report that in 2017 the aggregate value of "Individuell Pensjonssparing" (IPS) accounts was around 37 billion NOK, or 0.7% of total retirement wealth in Norway. Around 353 thousand people aged 17 and over had such retirement accounts, with a median balance of 48 thousand NOK. The vast majority of retirement wealth is

security-level ISIN numbers for publicly traded stocks, we merge the stock ownership data with prices and returns for all publicly traded stocks on the Oslo stock exchange (OSE). These returns account for stock splits and other similar events, allowing us to construct the buy-and-hold market return on the household's portfolio of Norwegian stocks.[10] We use the shareholder register information on shares in privately held companies to determine entrepreneurship, as discussed further below.

We obtain household balance sheet and income information from tax records ("Inntekt" register). The household balance sheet information includes total gross financial wealth subject to the Norwegian wealth tax and asset holdings for broad asset classes such as deposits, publicly-traded Norwegian stocks, stock and bond mutual funds, bonds, and foreign assets. It also comes from third-party reporting to the Norwegian tax authority (except ownership of foreign assets). We do not know the details of specific asset holdings within broad asset classes of financial wealth outside of publicly-traded Norwegian stocks, so the variation in portfolio returns will come only from the Norwegian stock component. However, being able to quantify the "known unknown" in financial wealth will prove important in our research design.

The Norwegian register data also provide a number of variables used as covariates in the analysis, including education and age of the household members, family status, and municipality of residence ("Befolkning" and "Utdanning" registers). We obtain the NACE sector of primary employment of the highest earning individual in the household ("household head") by merging the tax records to the employer-employee register ("Aa-registeret").

Our firm-level data start with information from the "Aksjonærregisteret" on all limited liability firms in Norway, including the exact foundation date, closing date (if the firm is dissolved), primary sector, the total number of shareholders for different classes of shares, and the book value of outstanding shares for each firm. We combine with information on employment from the employer-employee register, as well as annual firm balance sheets and income statements from tax records ("Regnskapsdata").[11] We use the employment in the subsequent year in the cases when employment in year of foundation is missing.

We restrict attention to households with a household head between 20 and 65 years old. We drop household-year observations with no earnings and zero financial wealth, as well as observations with (lagged) real gross financial wealth below 50 thousand and above 5 million Norwegian kroner. We also drop observations of direct stock owners with less than 1% (lagged) exposure to domestic publicly traded stocks. Furthermore, we drop from our

in the national social insurance fund, with a smaller amount managed by occupational pensions.

[10] See Ødegaard (2013) for details of the OSE data.

[11] We deflate all nominal values and returns to 2010 Norwegian kroner. Throughout our sample period the dollar-kroner exchange rate fluctuates between 4.9 and 9.3, with a mean of 6.8.

sample households after they become entrepreneurs. Our final sample covers the period 2004-2019.

3.2 Entrepreneur Definition

We define an entrepreneur as an individual who owns at least 1/3 of the book value of stocks in an incorporated non-financial firm, and where the firm has at most 3 stock owners and at least one employee in the year of foundation or subsequent year. We further require that either the newly-created firm does not own publicly traded domestic stocks in the year of foundation or that it has employees that are not members of the entrepreneur's household. A transition to entrepreneurship requires both a newly-created firm and that the household not have owned stocks in any private firm in the past. Upon transitioning to entrepreneurship, a household exits our sample.

These restrictions collectively focus attention on first-time active owners of new firms.[12] In particular, the limit on number of owners helps to exclude passive investment positions in private firms, the employment restriction ensures the new firm is economically active, and the requirement that either the firm have employees unrelated to the entrepreneur or that the firm holds no public equity helps to filter out inactive "family investment firms" created to store unrealized capital gains or losses for tax simplification purposes.

To put our entrepreneurship definition into perspective, Table 1 reports shares of business ownership and transitions to different types of business ownership both among the owners of publicly traded Norwegian stocks ("Stock owners"), as well as for our whole sample ("Population"). The first row includes both owners of at least 1/3 of the book value of any incorporated firm as well as households that receive non-incorporated business or farm income. We then progressively tighten the definition until we arrive at the definition of entrepreneurship we use.[13]

In addition to this summary statistics table, Appendix Table A.4 includes additional descriptive statistics for the groups of entrepreneurs and non-entrepreneurs, as well as the groups of direct owners of Norwegian public stocks and the rest. Entrepreneurs tend to be younger, with slightly higher earnings but slightly lower financial wealth and lower holdings of publicly traded domestic stocks compared to non-entrepreneurs. Direct owners of domestic

[12]See Brandt et al. (2022) for an analysis of the differences between serial and non-serial entrepreneurs using detailed firm-level data from China.

[13]Strictly speaking our baseline definition for entrepreneurs includes those in (VII) but for whom the transition happens for the first time. Therefore, in principle the definition in (VII) and our baseline definition of entrepreneurship do not overlap exactly because of the possibility of serial entrepreneurship with breaks. In practice, however, the difference is negligible as a share of the population, and so the shares in row (VII) essentially overlap with the shares of entrepreneurs according to our definition.

publicly traded stocks constitute around 12.6% of our sample. These households tend to be older, have higher earnings and higher level of financial wealth than the rest of our sample. The fact that entrepreneurs have higher earnings upon transitioning compared to non-entrepreneurs but lower direct holdings of domestic publicly traded stocks is broadly in line with the findings of Bhandari et al. (2022) using data from the U.S.

Table 1: Business Owners Descriptive Statistics

	Stock owner (in %)	Population (in %)
(I) owns $\geq \frac{1}{3}$ book value (BV) of an incorporated firm or receives business/farm/forestry income	19.11	17.75
(II) owns $\geq \frac{1}{3}$ BV of an incorporated firm ("owns a business")	6.07	4.98
(III) owns a business with $\leq$ 3 shareholders	4.95	4.22
(IV) AND is non-financial firm	4.39	3.83
(V) AND has employees	2.32	2.27
(VI) transitions to a business such as in (V)	0.87	0.82
(VII) transitions to a business such as in (V) that is newly created and not "family investment firm"	0.18	0.21

Notes: The table reports shares of business ownership and transitions to different types of business ownership both among the owners of publicly traded Norwegian stocks ("Stock owners") as well as for our whole sample ("Population"). The first row includes both owners of at least 1/3 of the book value of any incorporated firm as well as households that receive non-incorporated business or farm income. Subsequent rows progressively narrow this group as described in the first column of the table. Note that our baseline definition for entrepreneurs includes those in (VII) but for whom the transition happens for the first time. Therefore, in principle the definition in (VII) and our baseline definition of entrepreneurship do not overlap exactly because of serial entrepreneurship with breaks in the data. In practice, however, the difference is negligible as a share of the population and the shares in row (VII) essentially coincide with the shares of entrepreneurs as per our baseline definition.

4 Econometric Methodology

This section presents our baseline specification and explains how it addresses the main threats to causal identification.

4.1 Econometric Design

Let $E_{i,t}$ denote an indicator for individual i becoming an entrepreneur in year t. We model $E_{i,t}$ as a function of the return on financial wealth $r^*_{i,t}$, other *ex ante* observed characteristics $X_{i,t-1}$, and unobserved characteristics $\epsilon_{i,t}$.[14] The terms $X_{i,t-1}$ and $\epsilon_{i,t}$ include determinants

[14] Our model in Section 2 relates $E_{i,t}$ to assets rather than returns. These are equivalent if all individuals start at the same asset level. Otherwise, working in return space avoids a mechanical relationship between

discussed in Section 2 such as baseline financial wealth, the wage if remaining in paid employment, and entrepreneurial ability, as well as other factors such as preferences.

We do not observe $r^*_{i,t}$, because while we observe total financial wealth at the end of each year we do not observe all transactions during the year. Instead, we construct $r_{i,t}$ as the buy-and-hold return on the stock part of the wealth portfolio scaled by the share of initial wealth in stocks. Formally, we define $r_{i,t} = r^f_t + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r^f_t is the risk-free return, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. Below, we explain how we isolate quasi-random variation in $r_{i,t}$ by including a suitable array of fixed effects. Under this condition, a regression of the total wealth return $r^*_{i,t}$ (after removing contemporaneous buying or selling that may be correlated with $\epsilon_{i,t}$) on $r_{i,t}$ and the same fixed effects would yield a coefficient of one, reflecting the restriction to variation in $r_{i,t}$ from quasi-random portfolio choices uncorrelated with returns in the non-stock part of the portfolio. We therefore impose this "first stage" coefficient of one and directly model outcomes in terms of $r_{i,t}$.

Two main threats to causal identification remain. To frame them, it helps to decompose the stock portfolio excess return into systematic and idiosyncratic components: $\mathbf{s}'_{i,t-1}\mathbf{r}_t = \beta_{i,t-1} \times r^m_t + \nu_{i,t}$, where r^m_t is the market excess return in year t, $\beta_{i,t-1}$ is the portfolio "beta" for stock holdings at end of $t-1$, and $Var(\nu_{i,t}) = \sigma^2_{i,t-1}$.[15] The first threat arises because the *realized* idiosyncratic component $\nu_{i,t}$ may be correlated with unobserved determinants of entrepreneurship $\epsilon_{i,t}$. For example, home bias in portfolio choice (Coval and Moskowitz, 1999) could result in households experiencing better stock market returns in periods when their current industry or local area is booming. We address such concerns by including sector×time and municipality×time fixed effects in all specifications.

The second threat arises because *expected* returns may vary across households in a manner correlated with the entrepreneurship decision. For example, a more risk-tolerant individual might choose a stock portfolio with a higher market beta, implying higher expected returns, and risk tolerant individuals might also be more likely to transition to entrepreneurship for other reasons. Or individuals likely to become entrepreneurs might hold more or less of their wealth in domestic stocks. Borusyak and Hull (2021) term this "non-random exposure" and show that it suffices to control for the *ex ante* expected realization. In our setting, variation in exposure comes from portfolio characteristics. Specifically, the expected total

higher initial assets and being in extrema bins of changes in assets. In addition to interpreting the marginal effect of a higher return at the median asset level, we estimate a specification in level changes in robustness.

[15]We use this timing notation because we hold fixed the characteristics $\beta_{i,t-1}$ and $\sigma^2_{i,t-1}$ at their values from the previous year. We omit expected excess returns ("alpha") from the return representation because we find that the idiosyncratic component of returns has a small and *negative* serial correlation in our data.

buy-and-hold portfolio return is $\mathbb{E}_{t-1} r_{i,t} = \mathbb{E}_{t-1} r_t^f + \omega_{i,t-1} \times \beta_{i,t-1} \times \mathbb{E}_{t-1} r_t^m$.

We control flexibly for different expected returns by creating bins of $\omega_{i,t-1}, \beta_{i,t-1}$, and $\sigma_{i,t-1}$ and including interactions of these bins and time fixed effects, where the interactions with time accommodate unrestricted time-variation in the risk-free rate or expected market return, and the inclusion of $\sigma_{i,t-1}$ is necessary for non-linear or non-parametric specifications.[16] Effectively, we compare entrepreneurship rates across two individuals in the same year with the same allocation to domestic stocks and the same portfolio beta but different realized excess returns. The variation in $r_{i,t}$ thus comes from random realizations on portfolios with the same *ex ante* characteristics, where the randomness arises as the result of the idiosyncratic component of the portfolio holdings $\nu_{i,t}$ purged of industry or location characteristics.

To summarize, our baseline specification for the effect of the stock market on transitioning to entrepreneurship takes the form:

$$E_{i,t} = b \times r_{i,t} + \alpha_{\text{sector} \times t}(i) + \alpha_{\text{munic.} \times t}(i) + \alpha_{\beta \times \sigma \times \omega \times t}(i) + \epsilon_{i,t}, \tag{6}$$

where $\alpha_{y \times z}(i)$ denotes a fixed effect for observation i belonging to group $y \times z$. In some specifications we also control for additional covariates. These covariates absorb residual variation in the likelihood of becoming an entrepreneur and relax specific assumptions in our baseline implementation.

4.2 Implementation

We measure the portfolio characteristics $\beta_{i,t-1}$ and $\sigma_{i,t-1}$ using daily returns over the year $t-1$. Specifically, for each observation we form the time series of daily returns $\mathbf{s}_{i,t-1}\mathbf{r}_{t-1+\Delta}$, where $\mathbf{r}_{t-1+\Delta}$ gives the vector of individual stock returns on day $t-1+\Delta$, and we fix the weights at their value at the end of the year.[17] We obtain $\beta_{i,t-1}$ as the OLS regression coefficient from a regression of $\mathbf{s}_{i,t-1}\mathbf{r}_{t-1+\Delta}$ on $r^m_{t-1+\Delta}$ (which we equate with the OSE OBX stock market index) and $\sigma_{i,t-1}$ as the variance of $\mathbf{s}_{i,t-1}\mathbf{r}_{t-1+\Delta}$.[18] Figures A.1-A.3 in the Appendix plot the distribution of β, σ and ω.

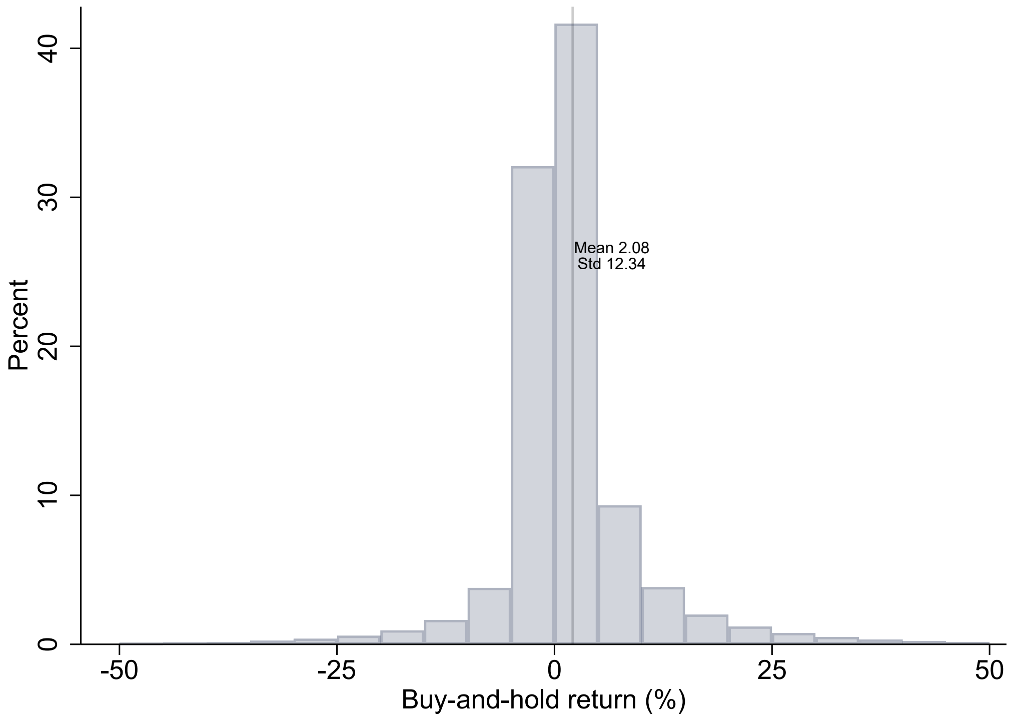
Figure 2 plots the unconditional distribution of (one year) buy-and-hold returns on fi-

[16] The bins for ω, β, and σ should have relatively low within-bin dispersion. We therefore truncate their distributions at the 2nd and 98th percentiles and report robustness in Section 5.2 to not truncating.

[17] We use price returns to focus on the unexpected component of the stock return but our results are little changed if we use total returns instead.

[18] Section 5.2 reports robustness to allowing for up to three portfolio factors or to replacing the OSE OBX with the U.S. CRSP value-weighted index, as well as several other robustness exercises that try to account for possible additional differences in portfolio characteristics that correlate with the propensity to start a firm.

Figure 2: Portfolio Return Distribution



Notes: Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t.

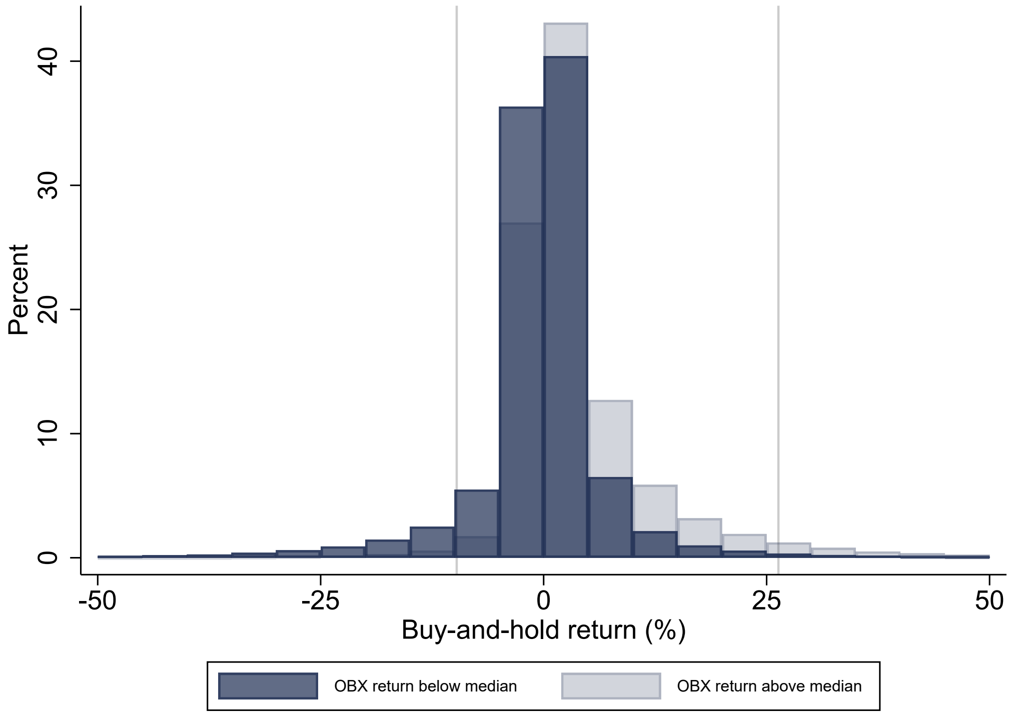
nancial wealth among the direct owners of publicly traded stocks in our sample.[19] Portfolio returns tend to be small on average but with a substantial standard deviation of around 12% and non-zero mass of relatively large return realizations of above 25%. The distribution of returns is also right-skewed, reflecting the in-sample positive mean return on the aggregate stock market, which we decompose in Figure 3 by splitting buy-and-hold portfolio returns in two groups based on whether the average return on the OSE OBX stock market index is above or below the 2004-2019 median.[20] In our empirical analysis below we will examine the heterogeneous effects of higher returns in years with above or below median aggregate stock market returns, which we will refer to as "good" and "bad" stock market years.

Where does variation in returns across households come from? Table A.2 reports statistics

[19]For households who are not direct owners, buy-and-hold returns will be set to zero, and these households will be treated as a separate category in our regressions.

[20]See also Figure A.6 for the distribution of buy-and-hold returns for each year in the period 2004-2019 and Figures A.4 and A.5 for the distribution of residualized returns conditional on the set of fixed-effects in our baseline specification.

Figure 3: Portfolio Return Distribution: Good vs. Bad Stock Market Years



Notes: Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. The figure plots the distributions of buy-and-hold returns for two groups of years: years in which the aggregate OBX return is above the median value in the period 2004-2019 and years in which the OBX return is below the median value.

from the distribution of household portfolio characteristics. Portfolios of domestic stocks exhibit high concentration, with the 90th percentile of number of stocks held just five. Such high concentration implies an absence of diversification, making possible sizable idiosyncratic differences in returns. Direct stock owners tend to have a relatively limited investment in other risky assets such as stock mutual funds with the remaining share of financial wealth held in deposits. Additionally, portfolios tend to be quite sticky, with a large share of households not making any portfolio adjustments over a one year horizon.

5 Results on the Propensity to Start a Business

In this section we present our main results on the effects of stock market wealth on the propensity to start a business.

5.1 Baseline Results

We start with a non-parametric approach. We partition the space of financial wealth and buy-and-hold portfolio returns into several bins and estimate the average effect from being in a particular bin on the propensity to start a business relative to a specific base bin. Our specification includes the fixed effect controls described in Section 4. We split financial wealth in two groups: below 600k ("moderate wealth") and above 600k ("high wealth"). For buy-and-hold returns we have 7 bins with the return bin of (-5%, 0%] serving as base.[21]

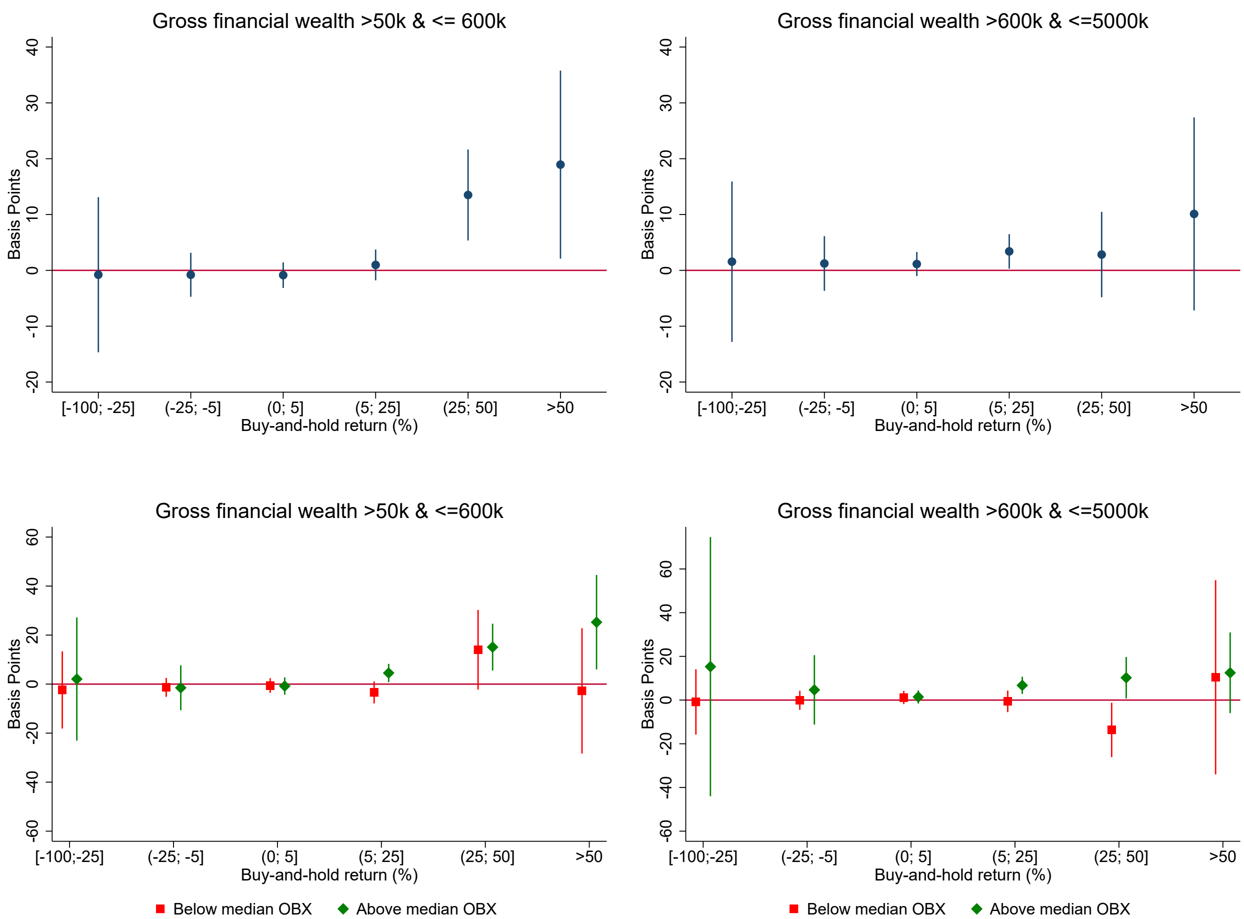
The first row of Figure 4 presents the estimated relative effects for each wealth bin. There is a notable positive effect of having a relatively high return of 25% or above for the moderate wealth group. In contrast the effects for high wealth are much smaller.

We also examine aggregate return heterogeneity by interacting each of the wealth-by-return bin with whether the aggregate stock return is above or below its median. The second row of Figure 4 shows the estimated relative effects (again with the (-5%, 0%] bin serving as base). There is a clear heterogeneity across good and bad stock market years. In bad stock market years the effect of returns is fairly flat and close to zero. In contrast, in good stock market years the propensity to start a business is increasing in return bins, particularly for the moderate wealth stock owners.

Table 2 reports regression coefficients. Column (1) pools the full sample and all years. The coefficient of 0.10 translates into a 2 basis point higher transition rate into entrepreneurship following a 20% stock return. Motivated by the evidence in Figure 4, column (2) restricts the sample to households with less than 600k, and column (3) allows the coefficient to vary between the moderate and high wealth groups. The coefficient estimate is larger for the sample excluding high wealth households and is also larger when comparing moderate to high wealth households. In particular, a two sided t-test rejects equality of the coefficients for moderate vs. high wealth households at a significance level of 5%. Furthermore, column (4) shows that the effects are significantly larger for good stock market years compared to bad stock market years among the moderate wealth group. A two-sided t-test rejects equality of the coefficient estimates for a good vs. bad stock market year in column (4) at a significance level of 1%. To put these magnitudes in context, the estimated effect in column (4) implies that a 20 p.p. increase in (one year) portfolio returns increases the propensity to start a business in a good stock market year by around 6.5 basis points, which is around 1/3 of the baseline entrepreneurship rate. Furthermore, the median financial wealth for the moderate wealth group of direct stock owners is close to 176k NOK, so a return of 20% corresponds to an increase in wealth for a household in the moderate wealth group with median financial

[21] We also include an additional bin for households who are not direct domestic stock owners. Alternative thresholds for the moderate wealth group around 600k NOK deliver similar estimated effects.

Figure 4: Non-parametric Entry Results



Notes: We define an entrepreneur as an individual that owns more than 1/3 of the book value of stocks in an incorporated non-financial firm with at most 3 stock owners which employs at least one worker. When considering the transition to entrepreneurship we only consider newly-created firms and households who have not owned stocks in any private firm in the past. In addition we require that the newly-created firm does not own publicly traded domestic stocks in the year of foundation unless it has employees that are not members of the entrepreneur's household. Furthermore, upon transitioning to entrepreneurship a household is dropped from our sample. Effects in each return bin are relative to a base buy-and-hold portfolio return between -5% and 0%. The second row shows effects in years when the OBX index return is above or below the median for the period 2004-2019. Controls include age group indicators (for 3 age groups), municipality-by-year fixed effects, primary employment sector-by-year fixed effects, and a set flexible controls given by a four-way interactions between interaction between 8 bins of exposure to directly-held domestic stocks, 7 bins of betas for the observed part of the portfolio, 7 bins of volatility for the observed part of the portfolio and year. Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. The bars represent 95% confidence intervals computed with clustering on the level of the municipality.

wealth of 35k NOK, or approximately 5k USD based on the sample average exchange rate.

Table 2: Baseline Entry Results

	Dep. var.: becomes entrepreneur ($E_{i,t}$)			
	(1)	(2)	(3)	(4)
$r_{i,t}$	0.109*	0.226**		
	(0.045)	(0.073)		
$r_{i,t}$, gross fin. wealth $> 600k$ NOK			0.023	
			(0.042)	
$r_{i,t}$, gross fin. wealth $\leq 600k$ NOK			0.170**	
			(0.061)	
$r_{i,t}$, $r_{OBX,t} \leq r_{OBX,median}$				−0.020
				(0.102)
$r_{i,t}$, $r_{OBX,t} > r_{OBX,median}$				0.321**
				(0.090)
Sample	All	$\leq 600k$	All	$\leq 600k$
Age group FE	Yes	Yes	Yes	Yes
Location-year FE	Yes	Yes	Yes	Yes
Primary sector-year FE	Yes	Yes	Yes	Yes
Portfolio-year FE	ω-β-σ	ω-β-σ	ω-β-σ	ω-β-σ
R^2	0.002	0.002	0.002	0.002
Clusters	422	422	422	422
Observations	15,447,959	11,708,791	15,447,959	11,708,791
Median fin. wealth	256.2	176	256.2	176
Median fin. wealth $\leq 600k$ NOK			176	
Median fin. wealth $> 600k$ NOK			1110.7	

Notes: The flexible controls in all specifications include a four-way interaction between 8 bins of exposure to directly-held domestic stocks, 7 bins of betas for the observed part of the stock portfolio, 7 bins of volatility for the observed part of the stock portfolio and year. Specifications (2) and (4) are restricted to gross financial wealth of up to 600k real NOK. We define an entrepreneur as an individual that owns more than 1/3 of the book value of stocks in an incorporated non-financial firm with at most 3 stock owners which employs at least one worker. When considering the transition to entrepreneurship we only consider newly-created firms and households who have not owned stocks in any private firm in the past. In addition we require that the newly-created firm does not own publicly traded domestic stocks in the year of foundation unless it has employees that are not members of the entrepreneur's household. Furthermore, upon transitioning to entrepreneurship a household is dropped from our sample. Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. All coefficient estimates are scaled by 100 for easier interpretation. Standard errors in parentheses are clustered at the municipality level. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

5.2 Robustness

We perform a number of robustness exercises and specification tests. In all of our robustness exercises we focus on the specification in column (2) of Table 2 as the baseline for concreteness. Table 3 presents robustness to changing the set of covariates. Column (1) adds 9 bins of financial wealth interacted with year to absorb any correlation between *ex ante* wealth and the propensity to start a firm. Column (2) controls for lagged log labor earnings, since as shown in Section 2 labor market earnings serve as an opportunity cost in the business creation decision of a potential entrepreneur. Column (3) includes a richer set of portfolio controls based on the Fama and French (1993) three factor model: a 6-way interaction between 8 bins of exposure to directly-held domestic stocks, 7 bins of market betas from a Fama-French 3-factor model, 7 bins of volatility for the observed part of the portfolio, 2 bins for loading on the SMB (small minus big) factor, 2 bins for loading on the HML (high minus low) factor and year. Column (4) controls for the interaction between the NACE sector of a direct owner's largest portfolio holding and year to account for stock-owners loading on industries that they believe would do well and subsequently starting firms in those industries. The main regression coefficient changes little with any of these additional covariates. Column (5) considers a specification that controls for portfolio composition by including a three-way interaction of bins of the share of financial wealth invested in domestic stock mutual funds and ETFs, bins of the exposure to directly-held domestic stocks and year. In this way we account for possible systematic differences in portfolio composition and the propensity to enter into business.[22] In particular, this flexible control can help account for systematic differences in financial literacy, which lead to households holding undiversified portfolios of stocks versus more diversified mutual fund holdings. These controls turn out to matter little for our coefficient estimate. Finally, column (6) illustrates the importance of the flexible controls for *ex ante* portfolio heterogeneity by removing them altogether. The estimated coefficient is now substantially reduced and less significant. This points to *ex ante* portfolio heterogeneity being an important confounder for the link between stock returns and entrepreneurship.

[22] In our measure of portfolio composition we focus only on stock mutual fund and domestic stock ownership, since other holdings of asset classes, such as bond mutual funds, bonds, or international financial assets are much more concentrated in Norway – see Table A.5.

Table 3: Robustness to Additional Covariates

	Dep. var.: becomes entrepreneur ($E_{i,t}$)					
	(1)	(2)	(3)	(4)	(5)	(6)
$r_{i,t}$	0.226**	0.212**	0.219**	0.218**	0.204*	0.083*
	(0.073)	(0.077)	(0.082)	(0.074)	(0.085)	(0.036)
Age group FE	Yes	Yes	Yes	Yes	Yes	Yes
Location-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Primary sector-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Financial wealth-year FE	Yes	No	No	No	No	No
Portfolio-year FE	ω-β-σ	ω-β-σ	3-factor model	ω-β-σ	ω-β-σ	None
R^2	0.002	0.002	0.003	0.002	0.002	0.002
Clusters	422	422	422	422	422	422
Observations	11,708,791	11,004,463	11,643,550	11,708,553	8,455,315	11,803,707
Description	Wealth bins	Labor income control	3-factor model	Addn'l. sectoral controls	Portfolio composition	No portfolio characteristics

Notes: Flexible controls in specifications (1), (2), (4) and (5) include a four-way interaction between 8 bins of exposure to directly-held domestic stocks, 7 bins of betas for the observed part of the portfolio, 7 bins of volatility for the observed part of the portfolio and year. Flexible controls in specification (3) include a 6-way interaction between 8 bins of exposure to directly-held domestic stocks, 7 bins of market betas from a Fama-French 3-factor model (Fama and French, 1993) for the observed part of the portfolio, 7 bins of volatility for the observed part of the portfolio, 2 bins for exposure to the SMB factor, 2 bins for exposure to the HML factor and year. Additional controls in specification (1) include 6 bins of financial wealth times year. Additional controls in specification (2) include lagged labor market income. Additional controls in specification (4) include an interaction between the level 1 NACE sector of the largest direct portfolio holding and year. Specification (5) includes a three-way interaction of 8 bins of share of financial wealth invested in directly-held domestic stocks, 5 bins of share of financial wealth held in domestic mutual funds and year. Note that for specification (5) we drop observations after 2015. We define an entrepreneur as an individual that owns more than 1/3 of the book value of stocks in an incorporated non-financial firm with at most 3 stock owners which employs at least one worker. When considering the transition to entrepreneurship we only consider newly-created firms and households who have not owned stocks in any private firm in the past. In addition we require that the newly-created firm does not own publicly traded domestic stocks in the year of foundation unless it has employees that are not members of the entrepreneur's household. Furthermore, upon transitioning to entrepreneurship a household is dropped from our sample. Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r^f_t + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r^f_t is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. All specifications are restricted to gross financial wealth of up to 600k real NOK. All coefficient estimates are scaled by 100 for easier interpretation. Standard errors in parentheses are clustered at the municipality level. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

Table 4: Robustness to Alternative Return Definitions

	Dep. var.: becomes entrepreneur ($E_{i,t}$)			
	(1)	(2)	(3)	(4)
$r_{i,t}$	0.242**	0.381*		
	(0.093)	(0.175)		
$r_{i,t}^{noemp}$			0.252**	
			(0.074)	
$r_{i,t+1}$				0.064
				(0.060)
Age group FE	Yes	Yes	Yes	Yes
Location-year FE	Yes	Yes	Yes	Yes
Primary sector-year FE	Yes	Yes	Yes	Yes
Financial wealth-year FE	No	No		
Portfolio-year FE	ω-β-σ	ω-β-σ	ω-β-σ	ω-β-σ
R^2	0.002	0.002	0.002	0.002
Clusters	422	422	422	422
Observations	11,521,512	11,296,110	11,708,791	11,655,703
Description	< 3 stock owner	< 3 stock owner (top 20 stocks)	Own employer stock	Placebo returns

Notes: Flexible controls in all specifications include a four-way interaction between 8 bins of exposure to directly-held domestic stocks, 7 bins of betas for the observed part of the portfolio, 7 bins of volatility for the observed part of the portfolio and year. We define an entrepreneur as an individual that owns more than 1/3 of the book value of stocks in an incorporated non-financial firm with at most 3 stock owners which employs at least one worker. When considering the transition to entrepreneurship we only consider newly-created firms and households who have not owned stocks in any private firm in the past. In addition we require that the newly-created firm does not own publicly traded domestic stocks in the year of foundation unless it has employees that are not members of the entrepreneur's household. Furthermore, upon transitioning to entrepreneurship a household is dropped from our sample. Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}_{i,t-1}'\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. $r_{i,t}^{noemp}$ is the buy-and-hold portfolio return that replaces the return of the firm which the head of the household is employed in with the OBX return for that year. The placebo return $r_{i,t+1} = r_{t+1}^f + \omega_{i,t-1}\mathbf{s}_{i,t-1}'\mathbf{r}_{t+1}$, where r_{t+1}^f is the risk-free return in year $t+1$, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_{t+1}$ the vector of realized excess returns of domestic stocks in year $t+1$. All specifications are restricted to gross financial wealth of up to 600k real NOK. All coefficient estimates are scaled by 100 for easier interpretation. Standard errors in parentheses are clustered at the municipality level. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

Table 4 explores changes to the return definition or restricting variation in returns. Column (1) narrows the stock portfolio variation to just households that hold directly less than 3 stocks. The coefficient estimate is largely unchanged, reflecting the fact that most of the variation in domestic stock portfolios in the data comes from owners of less than 3 stocks (see Table A.2). Column (2) further requires the main stock holding of such households to be among the 20 most popular companies traded by small investors on the Norwegian stock exchange, to help rule out undiversified investments in "exotic" single stocks due to private information or superior stock picking skill correlated with the propensity to start a firm.[23] In column (3) we account for stock owners holding stocks in their employer by replacing the firm-specific return for the firm where the household head is employed with the OBX return. This accounts for positive employer-specific returns due to innovative activity by the employer that may in turn trigger an idea spillover and spur additional business creation by employees (Babina and Howell, 2022; Chetty et al., in progress). Finally, column (4) performs a placebo exercise using the portfolio return in year $t+1$. The estimated coefficient is close to zero and insignificant in that case, bolstering the causal interpretation of our main effect.

Table A.7 in the Appendix reports results from additional robustness exercises. First, rather than computing the market beta using the domestic OBX index, which may be tilted towards energy stocks, we use the CRSP value-weighted index. Second, we consider the effects of trimming our sample for extreme realizations of exposure and stock portfolio characteristics. The estimates are slightly lower but still highly significant. Third, we restrict the sample only to households that do not receive business income in year $t-1$ to rule out possible changes in legal form of unincorporated businesses as opposed to new business creation. Fourth, in column (5) we modify the entrepreneurship definition by also considering existing firms that start hiring employees in addition to newly-created firms. We find positive and statistically significant effects of returns on entrepreneurship for this alternative definition as well. Fifth, we find a slightly smaller but still statistically significant response in a specification with buy-and-hold kroner gains or losses (in thousands of 2010 NOK) rather than portfolio returns, consistent with the marginal effect decreasing in wealth. Finally, to account for stock wealth held in non-taxable retirement accounts in the measure of gross financial wealth, we replace in the denominator of r the value of stock wealth held in Norwegian public stocks on tax returns with the value of the domestic stock portfolio from the stock register, with little change.

6 Results on Firm-level Outcomes

We now investigate how the entrepreneur's stock market wealth affects the characteristics of the new firm, using Proposition 3. In addition to being of interest in their own right, these results help to distinguish financial frictions from other explanations for the effect on firm entry.

[23]See Table A.6 in the Appendix for a list of these stocks.

6.1 Implementation Details

We implement the selection correction as follows. We restrict attention to non-negative stock returns and extend the exposition in Section 2 to incorporate $M = 4$ return bins: [0%,10%], (10%, 20%], (20%, 50%], and over 50%. Denote these return bins by $m \in \{1, 2, 3, 4\}$. For each bin below the highest, $m \in \{1, 2, 3\}$, denote by $p_m(x_i)$ the probability that the return is in the highest return bin rather than bin m, conditional on the covariates x_i. We estimate $p_m(x_i)$ using the sample of stock owners in bin m and in bin 4, controlling for the covariates from our baseline specification.[24] We obtain the propensity score weight $\omega(x_i)$ for each observation by applying the formula in Eq. (2), noting that $dF_a\left(a^H|x_i\right)/dF_a\left(a^m|x_i\right) = p_m(x_i)/(1 - p_m(x_i))$ and $dF_a\left(a^m\right)/dF_a\left(a^H\right)$ equals the relative share of the population in each bin.[25]

With the propensity score weights in hand, we calculate the reweighted probability $e^*(a^m, k \geq \bar{k})$ of becoming an entrepreneur for each return bin. Following Proposition 3, we then truncate the reweighted-distribution of initial assets of newly-started firms in return bin m at the $(1 - e_1^*/e_m^*) \times 100$-th percentile. This final step is the adjustment for selection into entrepreneurship. As discussed in **??**, truncating on initial assets helps to avoid post-entry fluctations in productivity that would break rank preservation. Finally, we estimate the effect of a higher portfolio return in the truncated sample via a weighted regression that uses the propensity score weights of firm income and balance-sheet statement and household financial outcomes on the average portfolio return in the entrepreneur's return bin. We report bootstrap standard errors that account for the estimated propensity score weights.

6.2 Baseline Results

Table 5 reports the results for firm income statement (top panel), balance sheet (middle panel) and household-level (bottom panel) outcomes in the year of foundation. The income and balance sheet items (except employment) are in thousands of 2010 NOK and annualized to adjust for differences in foundation dates. The household outcomes (apart from the pre-entry log earnings) are scaled by $t - 1$ gross financial wealth and multiplied by 100.

Higher wealth implies sizable positive effects for sales, employment, the wage bill, and value added. In terms of magnitudes, a 20 percentage point higher return increases these variables by between 30 and 60% of the mean. Crucially, higher wealth also increases total

[24] Due to computational costs associated with using a non-linear model on our large population and given the multiple fixed effects, we estimate these propensities using a linear probability model. Since the linear probability model may give propensities close to or above unity, which implies an undefined or very large value for the propensity score weight, we drop observations with estimated propensities above 90%.

[25] We additionally force the propensity score weights to average to unity, as advocated by Busso et al. (2009) to improve the performance of the propensity score reweighting procedure.

Table 5: Firm and Entrepreneur Outcomes

	Sales	Wage bill	Empl.	Value Added	VA/worker	EBITDA
$r_{i,t}$	54.7**	11.1*	0.04*	24.7*	−2.6	7.1**
	(19.5)	(5.4)	(0.02)	(11.3)	(4.2)	(2.7)
N	736	736	736	736	736	736
Mean	1832.8	662.9	2.2	1246.2	739.8	236.6
Median	1383	555.8	1	927.1	495.3	150.1

	Tot. assets	Fixed assets	Wk. Cap.	Equity	Tot. Liab.
$r_{i,t}$	21.2**	2.6*	3.5+	2.9*	17.6**
	(6.9)	(1.2)	(1.9)	(1.3)	(5.6)
N	736	736	736	736	736
Mean	797.5	119.8	176.9	200.4	546
Median	458	33.6	108.5	139.2	318.8

	Private firm equity	Change in stock holdings	Change in h.h. debt	Log of pre-entry earnings
$r_{i,t}$	1.10**	−0.23*	1.38	−0.003
	(0.35)	(0.09)	(1.28)	(0.01)
N	736	736	736	692
Mean	40.92	-1.96	28.4	13.07
Median	21.84	0	-8.93	13.23

Notes: All monetary values are in thousands of 2010 NOK. Outcomes are winsorized at the 5th and 95th percentiles. Working capital is defined as the difference between current assets and current liabilities. Entrepreneur balance sheet outcomes in the bottom panel are relative to lagged gross financial wealth. "Private firm equity" denotes the book value of a household's holdings of private firm equity relative to lagged gross financial wealth. "Change in stock holdings" is the change in the value of the portfolio of directly held stocks given constant stock prices between $t-1$ and t relative to lagged gross financial wealth. "Change in h.h. debt" is the change in total household debt over lagged gross financial wealth. "Log of pre-entry earnings" is the log of previous year's labor market earnings. The first three outcomes in the bottom table are scaled by 100 for easier interpretation. The results are based on the selection correction procedure described in Section 6. Bootstrapped standard errors in parentheses. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

earnings (EBITDA), consistent with a financial frictions channel but not a non-pecuniary benefit channel. In contrast, and consistent with our theoretical framework underpinning the selection correction procedure, there is a much smaller and insignificant effect on firm earnings without implementing the selection correction – see Table A.8 in the Appendix.

Turning to the balance sheet, total assets and fixed (tangible and intangible) assets also increase. The increase in owners' equity in the firm of 2.9k NOK for a 1% higher stock return implies a nearly one-to-one pass-through of stock wealth to owners' equity; specifically, a

1p.p. higher return on median financial wealth of 256k NOK amounts to 2.56k NOK higher stock wealth, almost exactly equal to the regression coefficient of 2.9.

The third panel provides evidence on how households finance the marginal increase in the size of their firm. The first three columns in this panel scale the dependent variable by lagged financial wealth, the same denominator as used to construct $r_{i,t}$, so that the coefficients have the interpretation of the marginal NOK change per additional NOK of stock wealth. Thus, household equity in the firm rises by 1.1 NOK for each additional 1 NOK of stock wealth and the data do not reject a pass-through of one. This pass-through mirrors the near one-to-one increase in total owners' equity in the firm in response to higher entrepreneur starting wealth. There is no mechanical reason that the estimated impact on owners' equity in the firm need coincide with the estimated impact on the household's holding of private firm equity, nor that either pass-through should lie near one. The fact that they do lends some credence to the selection correction procedure and bolsters the link between higher stock wealth and firm outcomes.

Households may fund their increase in private firm equity by liquidating publicly-traded stocks, borrowing, or using other savings. On average, households liquidate around 23 cents of their initial public equity position for every additional NOK of equity in the firm.[26] There is also a positive but statistically insignificant increase in household borrowing. The near one-to-one pass-through of marginal stock market wealth into firm equity and the evidence that households liquidate part of their portfolio and borrow to finance a larger firm further suggest the importance of liquidity constraints as a key friction making stock wealth relevant for business creation.

6.3 Robustness

Proposition 3 requires that the entry and asset choice decisions depend only on wealth a and business productivity z. Section 2.4 highlighted the robustness of these results to *ex ante* and *ex post* residual heterogeneity in size or profits which is independent from wealth a and business productivity z up to a potential bias close to the entry cutoff. In that case, looking at the subsample of firms which are sufficiently away from the entry cutoff mitigates the potential bias. Accordingly, Table A.9 in the Appendix presents a robustness exercise that accounts for residual heterogeneity by dropping the bottom 10% of firms in each return bin by size after the selection correction procedure. The coefficients change little, while the

[26] We also compare the average stock portfolio liquidation of entrepreneurs relative to non-entrepreneurs that are ex ante identical and who end up in the the same ex post return bin. For stock owners with portfolio return higher than 10%, entrepreneurs liquidate on average around 6% more of their stock holdings as a share of lagged financial wealth, compared to non-entrepreneurs.

standard errors increase due to the reduced sample size. Therefore, we conclude that our firm level outcome results appear robust to possible biases due to residual heterogeneity.

Finally, we include robustness exercises that address specific sources of residual heterogeneity. One concern is that individuals might vary in their outside option of labor income relative to business productivity. This heterogeneity can imply that some individuals with high labor income and relatively lower productivity enter only if they have higher initial wealth. This violation has a testable implication, since if the conditions in our model hold, then the distributions of z and the wage if work $w(z)$ should not vary with wealth after applying Proposition 3. Accordingly, the final regression reported in Table 5 has the pre-entry wage as the dependent variable. Consistent with the model's assumptions, we find no evidence of a difference in pre-entry wage by wealth.

A related concern is that individuals might have industry-specific skills that imply variation in their prospective firms' production processes or startup costs. Table A.3 reports the variation in median assets in the year of foundation by NACE sector. While most sectors have broadly similar initial sizes with assets in the range of 400-800k NOK, a few have much larger typical sizes, with utilities the largest at more than 14,000k NOK. Interestingly, Table A.3 also shows that sectors with very high typical initial size account for a relatively small share of new firms. This suggests that differences in capital intensity likely do not matter much for our firm-level outcomes. Table A.10 confirms this intuition by showing that the firm-level results are similar to and if anything for the most part slightly larger than our baseline when removing firms in sectors with high or very low start-up size.[27] The data do not reject equality of any coefficient reported in Table 5 with its counterpart in Table A.10.

7 Conclusion

In this paper we provide evidence that more stock market wealth causally increases business creation. The effects concentrate among moderate financial wealth individuals and in years when aggregate stock returns are high. This confluence points to a special role for stock market wealth in that stock wealth increases precisely when the returns to entrepreneurship are high.

Determining the effect of wealth on firm outcomes requires accounting for selection into entrepreneurship. Applying our model-motivated selection correction, we find that wealthier entrepreneurs start larger, more profitable firms. Together with the absence of marginal

[27]The high start-up size sectors are Electricity, gas, steam, and air conditioning supply, Water supply, Mining and quarrying, Financial and insurance activities, Agriculture, forestry and fishing, and Real estate activities. The two sectors with low start-up sizes are Other service activities and Administrative and support service activities.

effects in very high wealth households and the near one-to-one pass-through of marginal stock market wealth into firm equity, the positive effect on profits signifies financial frictions as a key mechanism for why wealth affects entrepreneurship.

Finally, our firm-level findings illustrate the importance of household wealth for business creation and growth. Initial firm size is a key determinant of long-run firm size and performance (Sedláček and Sterk (2017), Sterk et al. (2021)), raising the possibility of long-run effects of the stock market on economic growth via a business creation channel.

References

Adelino, M., Schoar, A., Severino, F., 2015. House prices, collateral, and self-employment. Journal of Financial Economics 117 (2), 288–306.

Aghion, P., Bolton, P., 1997. A theory of trickle-down growth and development. The review of economic studies 64 (2), 151–172.

Andersen, S., Nielsen, K. M., 2012. Ability or finances as constraints on entrepreneurship? evidence from survival rates in a natural experiment. The Review of Financial Studies 25 (12), 3684–3710.

Babina, T., Howell, S. T., 2022. Entrepreneurial spillovers from corporate r&d. Journal of Labor Economics.

Banerjee, A. V., Newman, A. F., 1993. Occupational choice and the process of development. Journal of political economy 101 (2), 274–298.

Bellon, A., Cookson, J. A., Gilje, E. P., Heimer, R. Z., 2021. Personal wealth, self-employment, and business ownership. The Review of Financial Studies 34 (8), 3935–3975.

Bermejo, V. J., Ferreira, M. A., Wolfenzon, D., Zambrana, R., 2022. Do cash windfalls affect entrepreneurship? evidence from the spanish christmas lottery: Evidence from the spanish christmas lottery. mimeo.

Bhandari, A., Kass, T., May, T., McGrattan, E., Schulz, E., 2022. On the nature of entrepreneurship.

Black, S. E., Strahan, P. E., 2002. Entrepreneurship and bank credit availability. The Journal of Finance 57 (6), 2807–2833.

Blanchflower, D. G., Oswald, A. J., 1998. What makes an entrepreneur? Journal of labor Economics 16 (1), 26–60.

Borusyak, K., Hull, P., 2021. Non-random exposure to exogenous shocks.

Brandt, L., Dai, R., Kambourov, G., Storesletten, K., Zhang, X., 2022. Serial entrepreneurship in china.

Buera, F. J., 2009. A dynamic model of entrepreneurship with borrowing constraints: theory and evidence. Annals of Finance 5, 443–464.

Buera, F. J., Kaboski, J. P., Shin, Y., 2015. Entrepreneurship and financial frictions: A macrodevelopment perspective. economics 7 (1), 409–436.

Buera, F. J., Shin, Y., 2013. Financial frictions and the persistence of history: A quantitative exploration. Journal of Political Economy 121 (2), 221–272.

Busso, M., DiNardo, J., McCrary, J., 2009. Finite sample properties of semiparametric estimators of average treatment effects. forthcoming in the Journal of Business and Economic Statistics.

Cagetti, M., De Nardi, M., 2006. Entrepreneurship, frictions, and wealth. Journal of political Economy 114 (5), 835–870.

Catherine, S., 2022. Keeping options open: What motivates entrepreneurs? Journal of Financial Economics 144 (1), 1–21.

Cesarini, D., Lindqvist, E., Notowidigdo, M. J., Östling, R., 2017. The effect of wealth on individual and household labor supply: evidence from swedish lotteries. American Economic Review 107 (12), 3917–3946.

Cespedes, J., Huang, X., Parra, C., 2021. More money, more options? the effect of cash windfalls on entrepreneurial activities in small businesses.

Chetty, R., Dossi, G., Smith, M., Van Reenen, J., Zidar, O., Zwick, E., in progress. America's missing entrepreneurs.

Chodorow-Reich, G., Nenov, P. T., Simsek, A., 2021. Stock market wealth and the real economy: A local labor market approach. American Economic Review 111 (5), 1613–1657.

Corradin, S., Popov, A., 2015. House prices, home equity borrowing, and entrepreneurship. The Review of Financial Studies 28 (8), 2399–2428.

Coval, J. D., Moskowitz, T. J., 1999. Home bias at home: Local equity preference in domestic portfolios. The Journal of Finance 54 (6), 2045–2073.

De Mel, S., McKenzie, D., Woodruff, C., 2008. Returns to capital in microenterprises: evidence from a field experiment. The quarterly journal of Economics 123 (4), 1329–1372.

Di Maggio, M., Kermani, A., Majlesi, K., 2020. Stock market returns and consumption. The Journal of Finance 75 (6), 3175–3219.

DiNardo, J., Fortin, N. M., Lemieux, T., 1996. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. Econometrica 64 (5), 1001–1044.

Evans, D., Leighton, L. S., 1989. Some empirical aspects of entrepreneurship. American Economic Review 79 (3), 519–35.

Evans, D. S., Jovanovic, B., 1989. An estimated model of entrepreneurial choice under liquidity constraints. Journal of political economy 97 (4), 808–827.

Fagereng, A., Guiso, L., Malacrino, D., Pistaferri, L., 2020. Heterogeneity and persistence in returns to wealth. Econometrica 88 (1), 115–170.

Fairlie, R. W., 1999. The absence of the african-american owned business: An analysis of the dynamics of self-employment. Journal of labor Economics 17 (1), 80–108.

Fairlie, R. W., Krashinsky, H. A., 2012. Liquidity constraints, household wealth, and entrepreneurship revisited. Review of Income and Wealth 58 (2), 279–306.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics 33 (1), 3–56.

Fracassi, C., Garmaise, M., Kogan, S., Natividad, G., 2013. How much does credit matter for entrepreneurial success in the united states? Available at SSRN.

Gentry, W. M., Hubbard, R. G., 2004. Entrepreneurship and household saving.

Golosov, M., Graber, M., Mogstad, M., Novgorodsky, D., 2021. How americans respond to idiosyncratic and exogenous changes in household wealth and unearned income. Tech. rep.

Hall, R. E., Woodward, S. E., 2010. The burden of the nondiversifiable risk of entrepreneurship. American Economic Review 100 (3), 1163–1194.

Hamilton, B. H., 2000. Does entrepreneurship pay? an empirical analysis of the returns to self-employment. Journal of Political economy 108 (3), 604–631.

Heckman, J. J., Ichimura, H., Todd, P., 1998. Matching as an econometric evaluation estimator. The review of economic studies 65 (2), 261–294.

Hirano, K., Imbens, G. W., Ridder, G., 2003. Efficient estimation of average treatment effects using the estimated propensity score. Econometrica 71 (4), 1161–1189.

Holtz-Eakin, D., Joulfaian, D., Rosen, H. S., 1994. Sticking it out: Entrepreneurial survival and liquidity constraints. Journal of Political economy 102 (1), 53–75.

Hurst, E., Lusardi, A., 2004. Liquidity constraints, household wealth, and entrepreneurship. Journal of political Economy 112 (2), 319–347.

Hurst, E. G., Pugsley, B. W., 2017. Wealth, Tastes, and Entrepreneurial Choice. University of Chicago Press, pp. 111–151.

Hvide, H. K., Møen, J., 2010. Lean and hungry or fat and content? entrepreneurs' wealth and start-up performance. Management Science 56 (8), 1242–1258.

Itskhoki, O., Moll, B., 2019. Optimal development policies with financial frictions. Econometrica 87 (1), 139–173.

Jensen, T. L., Leth-Petersen, S., Nanda, R., 2022. Financing constraints, home equity and selection into entrepreneurship. Journal of Financial Economics 145 (2), 318–337.

Kerr, S. P., Kerr, W. R., Nanda, R., 2022. House prices, home equity and entrepreneurship: Evidence from us census micro data. Journal of Monetary Economics 130, 103–119.

Kerr, W. R., Nanda, R., 2009. Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship. Journal of Financial Economics 94 (1), 124–149.

Kerr, W. R., Nanda, R., 2011. Financing constraints and entrepreneurship. Handbook of Research on Innovation and Entrepreneurship. Cheltenham: Elgar, 88–103.

Kihlstrom, R. E., Laffont, J.-J., 1979. A general equilibrium entrepreneurial theory of firm formation based on risk aversion. Journal of political economy 87 (4), 719–748.

Landier, A., Thesmar, D., 2008. Financial contracting with optimistic entrepreneurs. The Review of Financial Studies 22 (1), 117–150.

Leland, H. E., Pyle, D. H., 1977. Informational asymmetries, financial structure, and financial intermediation. The journal of Finance 32 (2), 371–387.

Lindh, T., Ohlsson, H., 1996. Self-employment and windfall gains: evidence from the swedish lottery. The Economic Journal 106 (439), 1515–1526.

McKenzie, D., 2017. Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. American Economic Review 107 (8), 2278–2307.

Moll, B., 2014. Productivity losses from financial frictions: Can self-financing undo capital misallocation? American Economic Review 104 (10), 3186–3221.

Moskowitz, T. J., Vissing-Jørgensen, A., 2002. The returns to entrepreneurial investment: A private equity premium puzzle? American Economic Review 92 (4), 745–778.

Nenov, P. T., 2017. Endogenous leverage and advantageous selection in credit markets. The Review of Financial Studies 30 (11), 3888–3920.

Ødegaard, B., 1990-2019. Obi financial data at bi norwegian business school. Tech. rep., BI Norwegian Business School.

Ødegaard, B. A., 2013. Empirics of the oslo stock exchange. basic, descriptive, results 1980-2020. Working Paper.

Parker, S. C., 2018. The economics of entrepreneurship. Cambridge University Press.

Piketty, T., 1997. The dynamics of the wealth distribution and the interest rate with credit rationing. The Review of Economic Studies 64 (2), 173–189.

Pugsley, B. W., Hurst, E., 2011. What do small businesses do? Brookings Papers on Economic Activity 43 (2), 73–142.

Quadrini, V., 1999. The importance of entrepreneurship for wealth concentration and mobility. Review of income and Wealth 45 (1), 1–19.

Ring, M. A. K., 2022. Entrepreneurial wealth and employment: Tracing out the effects of a stock market crash. Available at SSRN 3716347.

Rosenbaum, P. R., Rubin, D. B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70 (1), 41–55.

Ross, S. A., 1977. The determination of financial structure: the incentive-signalling approach. The bell journal of economics, 23–40.

Schmalz, M. C., Sraer, D. A., Thesmar, D., 2017. Housing collateral and entrepreneurship. The Journal of Finance 72 (1), 99–132.

Sedláček, P., Sterk, V., 2017. The growth potential of startups over the business cycle. American Economic Review 107 (10), 3182–3210.

Smith, M., Yagan, D., Zidar, O., Zwick, E., 07 2019. Capitalists in the Twenty-First Century*. The Quarterly Journal of Economics 134 (4), 1675–1745.

Smogeli, P. O., Halvorsen, E., 2019. Pensjonsformue i Norge 2017. Statistisk sentralbyrå.

Statistics Norway, 2004-2014. Aa-registeret. Tech. rep., Statistics Norway.

Statistics Norway, 2004-2019a. Aksjonærregisteret. Tech. rep., Statistics Norway.

Statistics Norway, 2004-2019b. Education register ("utdanning"). Tech. rep., Statistics Norway.

Statistics Norway, 2004-2019c. Income register ("inntekt"). Tech. rep., Statistics Norway.

Statistics Norway, 2004-2019d. Population register ("befolkning"). Tech. rep., Statistics Norway.

Statistics Norway, 2004-2019e. Regnskapsdata. Tech. rep., Statistics Norway.

Statistics Norway, 2015-2019. A-ordningen. Tech. rep., Statistics Norway.

Sterk, V., Sedláček, P., Pugsley, B., 2021. The nature of firm growth. American Economic Review 111 (2), 547–79.

Stock Market Wealth and Entrepreneurship

Online Appendix

Gabriel Chodorow-Reich Plamen Nenov Vitor Santos Alp Simsek

A Data Appendix

A.1 Summary statistics

Table A.1: Summary Statistics of New Firms (thousands of 2010 NOK)

	All years		**Low return years**		**High return years**	
	Mean	Median	Mean	Median	Mean	Median
Total Assets	928.0	478.8	861.0	454.9	996.9	503.6
Sales	1851.5	819.0	1714.2	789.3	1992.9	848.7
Wage Bill	542.1	316.0	525.8	306.2	558.9	323.9
Employment	3.3	2.0	3.1	2.0	3.5	2.0
Value added	985.9	579.9	962.5	567.3	1010.0	593.0
Value added / worker	417.8	276.8	428.0	279.0	407.4	275.1
Working Capital	120.8	66.7	112.7	68.0	129.2	64.9
Fixed Assets	273.0	48.0	251.3	44.3	295.4	51.1
Tangibles	209.3	26.3	194.1	23.5	225.0	29.5
Equity	186.8	104.7	181.8	102.0	192.0	107.2

Table A.2: Distribution of Portfolio Characteristics for Stock Holders

Percentile	β	ω	σ	**Herfindahl index**	**Number of stocks**	**Change in stock holdings (in %)**	**Share stock mutual funds**	**Share deposits**
10th	0.45	0.02	0.01	0.34	1	-100.00	0.00	0.16
20th	0.63	0.03	0.01	0.47	1	-18.01	0.00	0.29
30th	0.76	0.05	0.01	0.56	1	-0.00	0.00	0.40
40th	0.87	0.07	0.02	0.76	1	0.00	0.00	0.50
50th	0.97	0.10	0.02	0.99	2	0.00	0.02	0.60
60th	1.00	0.14	0.02	1.00	2	0.00	0.06	0.69
70th	1.04	0.20	0.02	1.00	3	4.67	0.10	0.77
80th	1.16	0.29	0.03	1.00	3	25.37	0.17	0.85
90th	1.32	0.44	0.03	1.00	5	76.78	0.30	0.92

Notes: ω is the share of directly held domestic stocks out of total financial wealth. "Share stock mutual funds" is the portfolio share invested in domestic stock mutual funds and exchange-traded funds. "Share deposits" is the share of financial wealth held in deposits. Note that the stock mutual fund shares are for the period 2004-2015.

Table A.3: Industry Distribution of New Firms, 2009-2019.

Industry	Share (%)	Median total assets (1000 NOK)
Construction	23.10	467
Wholesale and retail trade and repair of motor vehicles	20.57	774
Professional, scientific and technical activities	16.08	398
Administrative and support service activities	6.71	376
Information and communication	5.93	396
Human health and social work activities	5.66	473
Manufacturing	4.49	757
Transportation and storage	4.38	661
Other service activities	3.59	293
Real estate activities	3.05	1297
Education	1.9	278
Arts, entertainment and recreation	1.81	394
Agriculture, forestry and fishing	1.35	1074
Financial and insurance activities	0.72	3572
Mining and quarrying	0.27	870
Water supply, sewerage, waste management	0.22	1380
Electricity, gas, steam and air conditioning supply	0.16	14133

Notes: Monetary amounts are in 1000s of 2010 Norwegian kroner.

Table A.4: Sample Descriptive Statistics.

	Non-Entrepreneur					Entrepreneur				
	Mean	**p25**	**p50**	**p75**	**Share**	**Mean**	**p25**	**p50**	**p75**	**Share**
Age	44.7	35	45	55	.	39.8	32	39	47	.
Earnings	648.7	339.2	591.9	912.9	.	655.3	362.7	608.1	875.3	.
Financial Wealth	579.1	120.0	269.7	639.9	.	528.1	143.2	291.5	606.2	.
Directly held stocks	143.2	13.4	46.6	137.8	.	108.9	6.4	29.3	98.3	.
Share	–	–	–	–	**99.79**	–	–	–	–	**.21**
	(Direct) Stock owner					**Other**				
	Mean	**p25**	**p50**	**p75**	**Mean**	**p25**	**p50**	**p75**	**Share**	
Age	47	38	48	57	.	44.3	34	45	55	.
Earnings	884.9	495.5	836.2	1176.4	.	614.8	322.1	560.7	874.7	.
Financial Wealth	1031.2	276.3	601.1	1298.0	.	514.0	110.5	241.3	557.9	.
Directly held stocks	154.5	18.6	54.4	151.2	.	0.0	0.0	0.0	0.0	.
Share	–	–	–	–	**12.56**	–	–	–	–	**87.44**
	Moderate Wealth (<600K)					**High Wealth (>600K)**				
	Mean	**p25**	**p50**	**p75**	**Share**	**Mean**	**p25**	**p50**	**p75**	**Share**
Age	43	33	43	53	.	49.6	42	51	59	.
Earnings	607.1	328.6	561.2	864.8	.	774.8	377.4	709.5	1079.4	.
Financial Wealth	270.6	96.0	187.1	345.7	.	1514.6	745.0	1117.6	1857.2	.
Directly held stock	59.2	7.3	24.9	69.3	.	244.4	34.1	98.2	271.3	.
Share	–	–	–	–	**75.21**	–	–	–	–	**24.79**

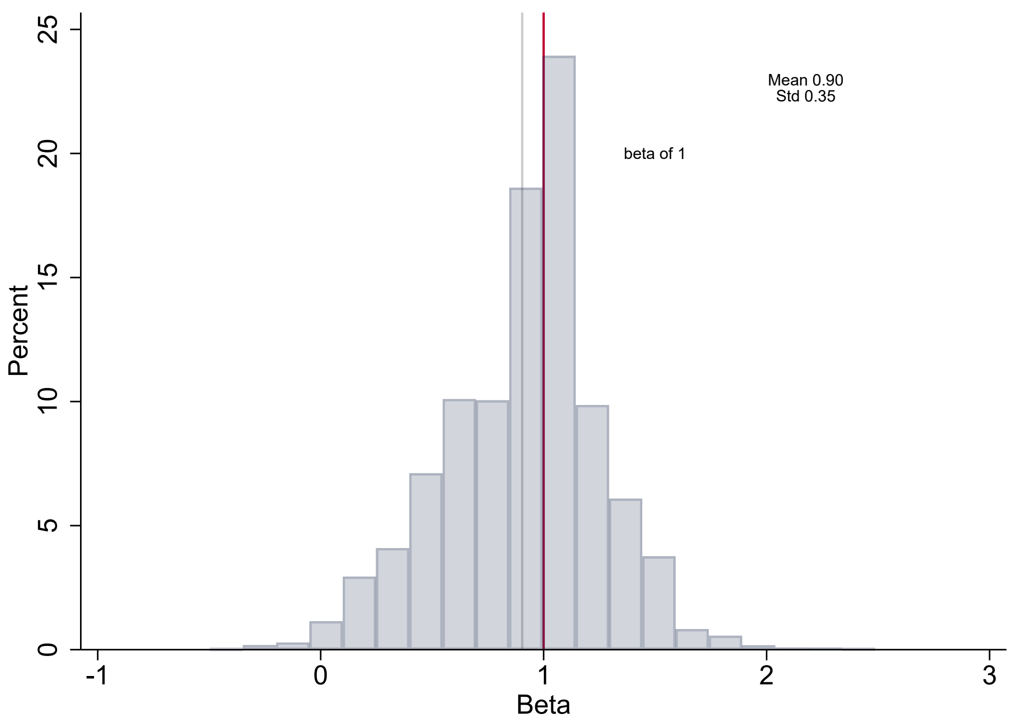
Notes: Monetary amounts are in 1000s of 2010 Norwegian kroner.

Table A.5: Shares of the Population by Asset Holdings (2004-2015).

	Population	Moderate Wealth ($\leq 600K$)	High wealth ($> 600K$)
1. Direct owner domestic stock	0.131	0.095	0.250
Of which: < 3 stocks	0.085	0.070	0.134
Of which: among top 20	0.055	0.044	0.093
2. Domestic stock mutual fund owner	0.409	0.366	0.550
3. Direct or mutual fund owner	0.459	0.409	0.621
4. Direct and mutual fund owner	0.082	0.051	0.180
Of which: < 3 stocks	0.050	0.037	0.093
Of which: among top 20	0.033	0.023	0.066
5. Domestic bond mutual fund owner	0.092	0.079	0.131
6. Domestic bond owner	0.012	0.006	0.032
7. Foreign assets owner	0.055	0.036	0.117

A.2 Distributions

Figure A.1: Beta of Portfolio of Directly Held Domestic Stocks



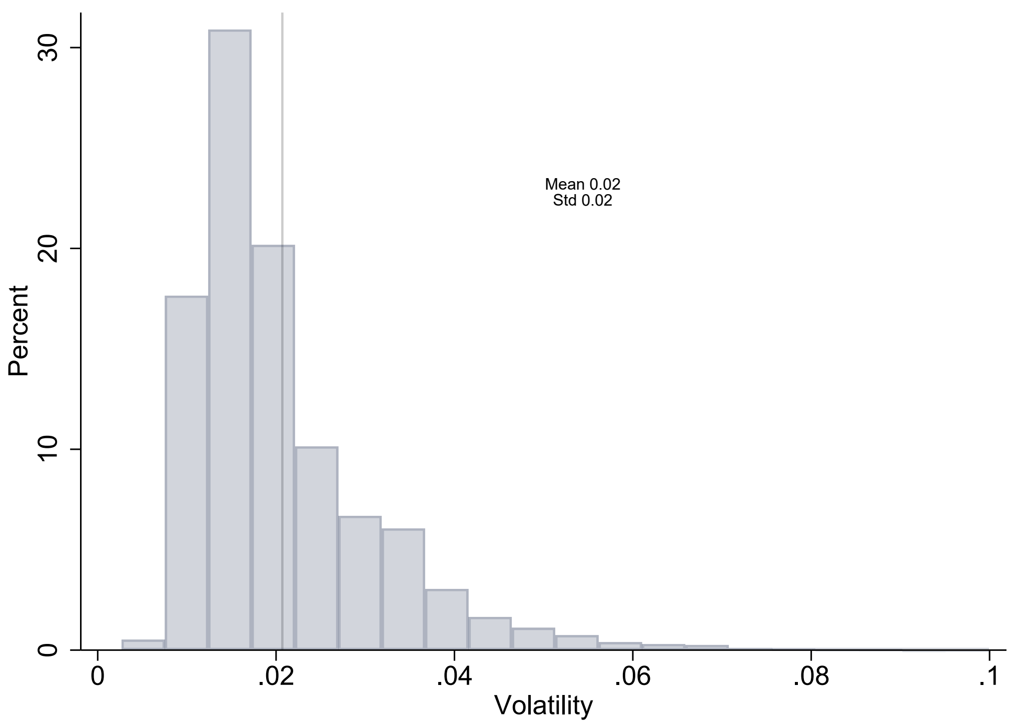
Notes: We measure the portfolio characteristics $\beta_{i,t-1}$ and $\sigma_{i,t-1}$ using daily returns over the year $t-1$. Specifically, for each observation we form the time series of daily returns $\tilde{r}_{i,t-1+\Delta} = \omega_{i,t-1}^{-1}\mathbf{s}_{i,t-1}\mathbf{r}_{t-1+\Delta}$, where $\tilde{r}_{i,t-1+\Delta}$ gives the return on day $t-1+\Delta$ of a portfolio with weights fixed at their value at the end of the year. We obtain $\beta_{i,t-1}$ as the OLS regression coefficient from a regression of $\tilde{r}_{i,t-1+\Delta}$ on $r^m_{t-1+\Delta}$ less the risk-free rate and $\sigma_{i,t}$ as the variance of $\tilde{r}_{i,t-1+\Delta}$.

Table A.6: List of the 20 Most Popular Publicly Traded Companies Held by Direct Owners of Less Than 3 Domestic Stocks

5-digit NACE sector	**Average Mkt. Cap.**	**Average Number**	**Fraction shareholders**	**Ownership share**	**Share wealth invested**
Extraction of crude petroleum	1083.2	36.8	0.35	0.01	0.21
Wireless telecom. activities	1053.0	18.9	0.35	0.01	0.07
Manufact. prepared meals and dishes	143.0	15.8	0.32	0.02	0.10
Non-life insurance	109.4	13.9	0.31	0.02	0.06
Activities financial holding companies	89.8	12.5	0.38	0.01	0.01
Manufacturing paper and paperboard	63.6	10.8	0.38	0.05	0.01
Production primary aluminium	547.7	9.0	0.15	0.01	0.04
Other technical consultancy	24.4	7.1	0.19	0.04	0.02
Construction oil-platforms and modules	554.9	5.2	0.28	0.01	0.01
Manufacture fertilisers & nitrogen compounds	120.0	5.2	0.11	0.003	0.01
Operation marine fish farms	152.6	5.0	0.18	0.01	0.02
Extraction of crude petroleum	3.6	4.5	0.20	0.05	0.03
Scheduled long-dist. bus transports	2.5	3.9	0.59	0.09	0.004
Scheduled long-dist. transport coastal waters	4.3	3.6	0.35	0.04	0.004
Construction residential & non-residential buildings	6.2	3.4	0.38	0.10	0.05
Other monetary intermediation	268.9	3.2	0.22	0.07	0.04
Electricity production through water power	14.5	3.1	0.28	0.01	0.01
Wholesale computers computer equip. and software	37.1	3.1	0.25	0.02	0.004
Freezing of fish, fish fillets crustaceans and molluscs	0.04	3.0	0.20	0.02	0.0004
Other monetary intermediation	172.8	2.5	0.37	0.14	0.02

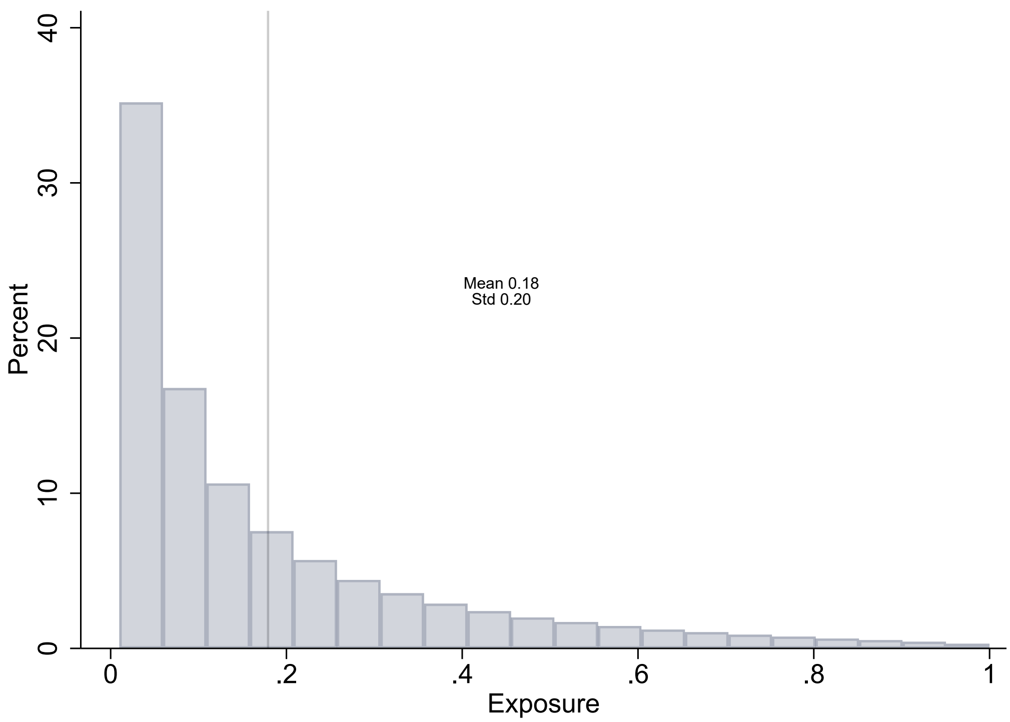
Notes: "Average Mkt. Cap" is the average market capitalization of the company during our sample period (in billions of NOK). "Average number" is the average number of owners of less than 3 stocks (in thousands). "Fraction shareholders" is the share out of all stockholders in the company who are owners of less than 3 stocks. "Ownership share" is the share of the firm owned by owners of less than 3 stocks. "Share wealth invested" is the average share of the total stock market wealth of owners of less than 3 stocks that is invested in that particular company.

Figure A.2: Volatility of Portfolio of Directly Held Domestic Stocks



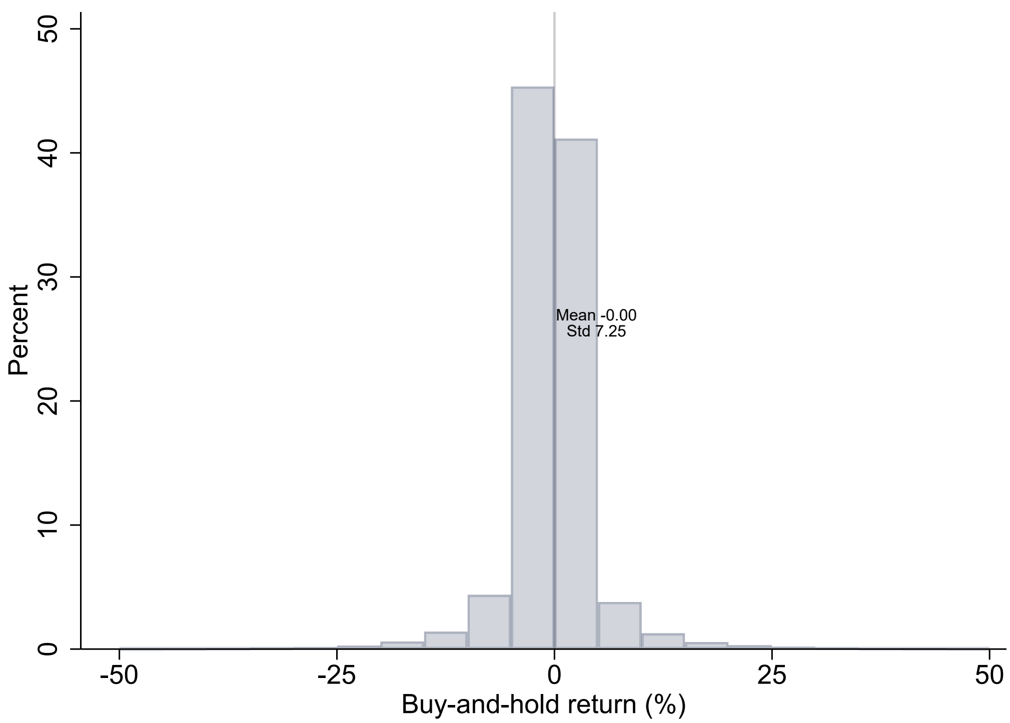
Notes: We measure the portfolio characteristics $\beta_{i,t-1}$ and $\sigma_{i,t-1}$ using daily returns over the year $t-1$. Specifically, for each observation we form the time series of daily returns $\tilde{r}_{i,t-1+\Delta} = \omega_{i,t-1}^{-1}\mathbf{s}_{i,t-1}\mathbf{r}_{t-1+\Delta}$, where $\tilde{r}_{i,t-1+\Delta}$ gives the return on day $t-1+\Delta$ of a portfolio with weights fixed at their value at the end of the year. We obtain $\beta_{i,t-1}$ as the OLS regression coefficient from a regression of $\tilde{r}_{i,t-1+\Delta}$ on $r^m_{t-1+\Delta}$ less the risk-free rate and $\sigma_{i,t}$ as the variance of $\tilde{r}_{i,t-1+\Delta}$.

Figure A.3: Exposure to Domestic Stocks



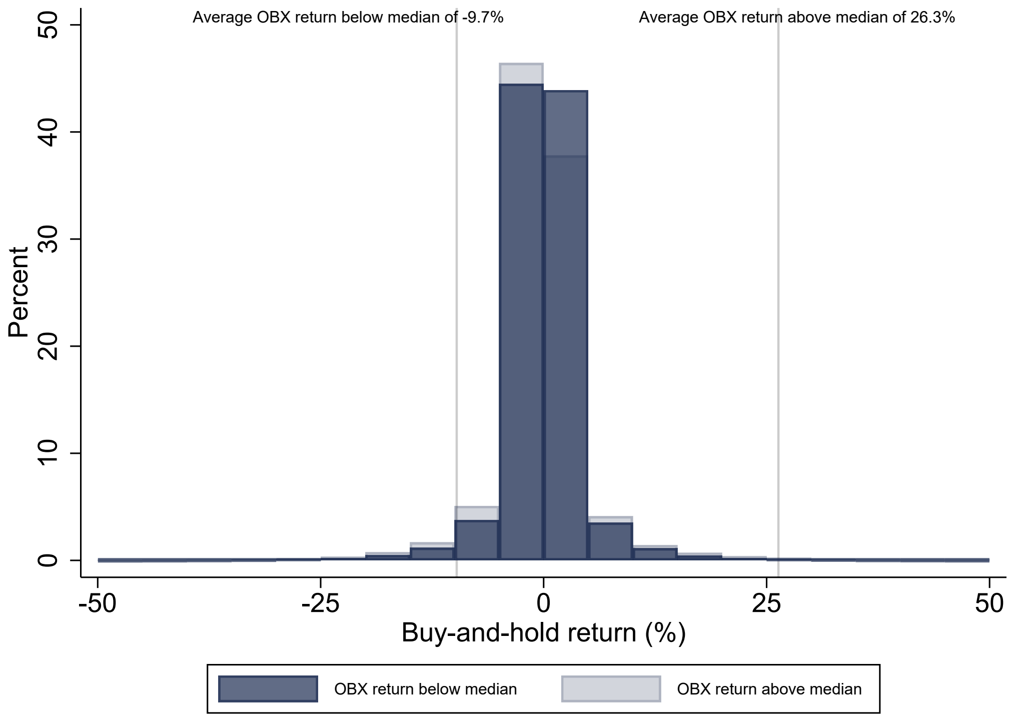
Notes: $\omega_{i,t}$ denotes the share of financial wealth held in domestic stocks at the end of year t.

Figure A.4: Portfolio Return Distribution (residualized)



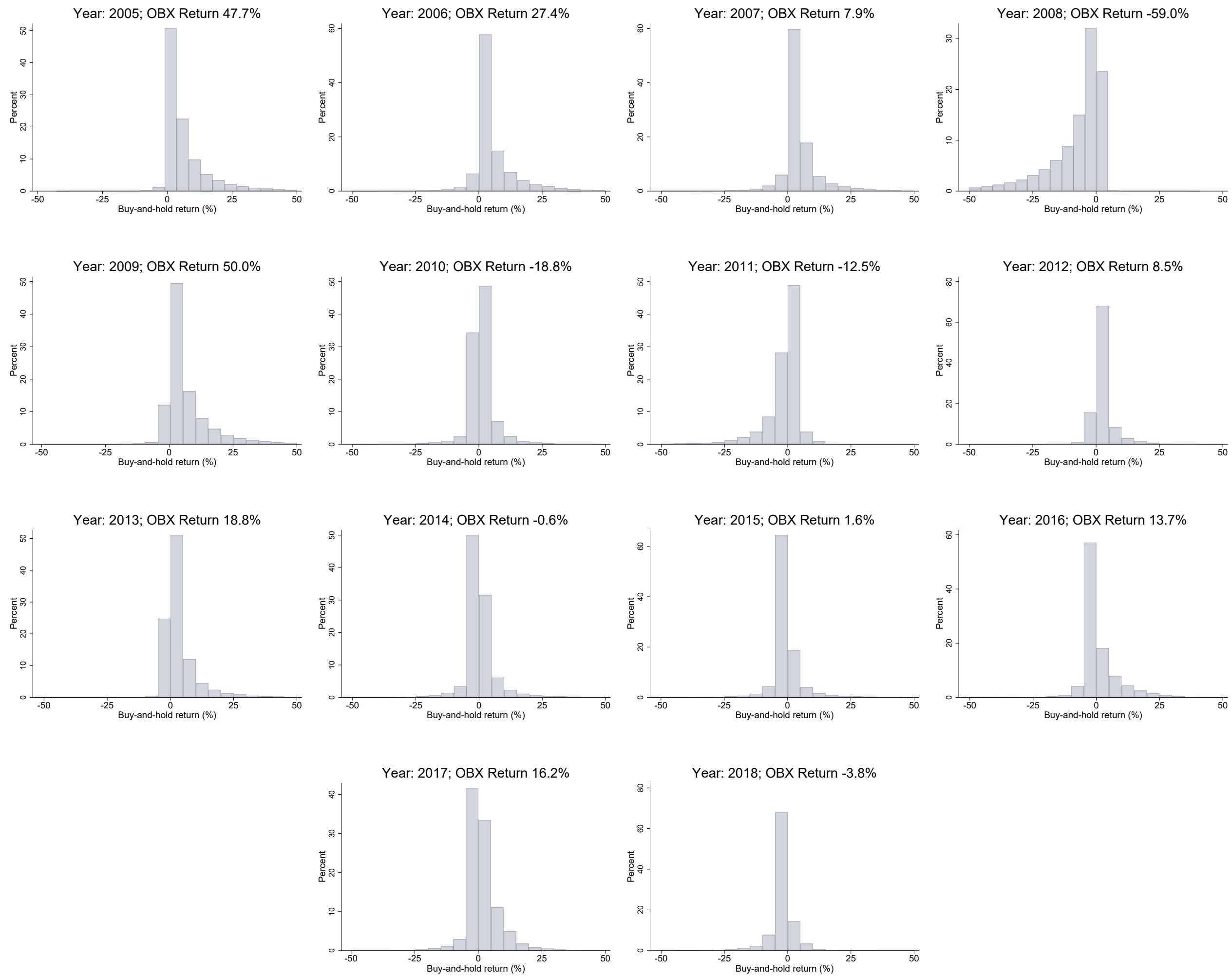
Notes: Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. The figure plots the distribution of residualized buy-and-hold portfolio returns after partialling out the fixed effects from the baseline specification.

Figure A.5: Portfolio Return Distribution: Good vs. Bad Stock Market Years (residualized)



Notes: Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. The figure plots the distributions of buy-and-hold portfolio returns for two groups of years: years in which the aggregate OBX return is above the median value in the period 2004-2019 and years in which the OBX return is below the median value. The figure plots the distribution of residualized buy-and-hold portfolio returns after partialling out the fixed effects from the baseline specification.

Figure A.6: Portfolio Returns by Year



Notes: Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t.

A.3 Additional Empirical Results

Table A.7: Additional Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
$r_{i,t}$	0.237**	0.145**	0.138*	0.248*		
	(0.078)	(0.049)	(0.065)	(0.116)		
pot. gain					0.063**	
					(0.020)	
$r_{i,t}^{alt}$						0.203**
						(0.066)
Age group FE	Yes	Yes	Yes	Yes	Yes	Yes
Loc.-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio-year FE	ω-β_{CRSP}-σ	ω-β-σ	ω-β-σ	ω-β-σ	ω-β-σ	ω-β-σ
R^2	0.002	0.002	0.002	0.003	0.002	0.002
Clusters	422	422	422	422	422	422
Observations	11,645,434	11,803,582	9,994,039	11,451,195	11,70,791	11,727,986
Description	CRSP index	No trimming	No business income	$E_{i,t}$ or starting employment	Pot. change in wealth	Alternative fin. wealth

Notes: Flexible controls include a four-way interaction between 8 bins of exposure to directly-held domestic stocks, 7 bins of betas for the observed part of the portfolio, 7 bins of volatility for the observed part of the portfolio and year. For specification (1) the betas are replaced with betas with respect to the CRSP value-weighted index. For specification (3) we restrict the sample to households that have not received business income in the year prior to the firm's foundation. For specification (5) we consider either our baseline entrepreneurship definition or existing firms transitioning to positive employment. We define an entrepreneur as an individual that owns more than 1/3 of the book value of stocks in an incorporated non-financial firm with at most 3 stock owners which employs at least one worker. When considering the transition to entrepreneurship we only consider newly-created firms and households who have not owned stocks in any private firm in the past. In addition we require that the newly-created firm does not own publicly traded domestic stocks in the year of foundation unless it has employees that are not members of the entrepreneur's household. Furthermore, upon transitioning to entrepreneurship a household is dropped from our sample. Buy-and-hold returns $r_{i,t}$ are defined as $r_{i,t} = r_t^f + \omega_{i,t-1}\mathbf{s}'_{i,t-1}\mathbf{r}_t$, where r_t^f is the risk-free return in year t, $\omega_{i,t-1}$ denotes the share of financial wealth held in domestic stocks at the end of year $t-1$, $\mathbf{s}_{i,t-1}$ the weight in the stock portfolio of each domestic stock, and $\mathbf{r}_t$ the vector of realized excess returns of domestic stocks in year t. Kroner return$_{i,t}$ denotes the buy-and-hold gain/loss in thousands of 2010 NOK. $r_{i,t}^{alt}$ denotes an alternative buy-and-hold portfolio return formed by constructing domestic stock exposure using an alternative measure of gross financial wealth formed by replacing gross financial wealth with an alternative measure that replaces the value of stock wealth held in Norwegian public stocks on the tax returns with the value of the domestic stock portfolio we compute directly. All specifications are restricted to gross financial wealth of up to 600k real NOK. All coefficient estimates are scaled by 100 for easier interpretation. Standard errors in parentheses are clustered at the municipality level. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

Table A.8: Firm and Entrepreneur Outcomes – No Selection Correction

	Sales	Wage bill	Empl.	Value Added	VA/worker	EBITDA
$r_{i,t}$	22.9*	5.5+	0.02+	9.1	−3.9	2.3
	(10.8)	(2.9)	(0.01)	(6.4)	(3.5)	(1.8)
N	797	797	797	797	797	797
Mean	1684.7	664.8	2.2	1146.2	682.1	202.5
Median	1330.1	591.1	1	909.8	427.9	97.1

	Tot. assets	Fixed assets	Wk. Cap.	Equity	Tot. Liab.
$r_{i,t}$	9.1*	1.3+	1.2	1.3	7.8*
	(4.1)	(0.8)	(1.4)	(0.9)	(3.2)
N	797	797	797	797	797
Mean	748.6	116.4	161.3	189.9	514
Median	458	42.5	60.2	139.2	318.8

	Private firm equity	Change in stock holdings	Change in h.h. debt	Log of pre-entry earnings
$r_{i,t}$	0.63**	−0.15+	1.23	0.002
	(0.22)	(0.08)	(0.92)	(0.005)
N	797	797	797	749
Mean	38.41	-2.84	35.77	13.13
Median	21.84	0	-1.84	13.31

Notes: The table corresponds to Table 5 without the selection correction procedure described in Section 6. All monetary values are in thousands of 2010 NOK. Outcomes are winsorized at the 5th and 95th percentiles. Working capital is defined as the difference between current assets and current liabilities. Entrepreneur balance sheet outcomes in the bottom panel are relative to lagged gross financial wealth. "Private firm equity" denotes the book value of a household's holdings of private firm equity relative to lagged gross financial wealth. "Change in stock holdings" is the change in the value of the portfolio of directly held stocks given constant stock prices between $t-1$ and t relative to lagged gross financial wealth. "Change in h.h. debt" is the change in total household debt over lagged gross financial wealth. The first three outcomes in the bottom table are scaled by 100 for easier interpretation. Bootstrapped standard errors in parentheses. The rows labeled P(=Table 5) report the p-values from a bootstrapped t-test of equality between the coefficients in this table and in Table 5. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

Table A.9: Firm and Entrepreneur Outcomes – 10% trim

	Sales	Wage bill	Empl.	Value Added	VA/worker	EBITDA
$r_{i,t}$	56.1*	11.1	0.05^{+}	23.2	−6.1^{+}	6.6^{+}
	(27.4)	(8.4)	(0.03)	(16.2)	(3.6)	(3.5)
N	662	662	662	662	662	662
Mean	2010	718.3	2.2	1375.6	808.9	265.8
Median	1502.6	607.3	1	1090.8	633.9	150.1

	Tot. assets	Fixed assets	Wk. Cap.	Equity	Tot. Liab.
$r_{i,t}$	21.5*	2.7	3.1	2.5	18.3^{+}
	(9.5)	(1.7)	(2.7)	(1.8)	(7.9)
N	662	662	662	662	662
Mean	881.5	132.5	196.6	221.8	603.9
Median	458	62.6	112.9	139.2	318.8

	Private firm equity	Change in stock holdings	Change in h.h. debt	Log of pre-entry earnings
$r_{i,t}$	1.00	−0.23	1.79	−0.003
	(0.70)	(0.16)	(1.5)	(0.01)
N	662	662	662	620
Mean	44.2	-.83	23.37	13.04
Median	21.84	0	-8.93	13.24

Notes: The table corresponds to Table 5 but dropping 10% of the lowest ranked firms by assets in each return bin. All monetary values are in thousands of 2010 NOK. Outcomes are winsorized at the 5th and 95th percentiles. Working capital is defined as the difference between current assets and current liabilities. Entrepreneur balance sheet outcomes in the bottom panel are relative to lagged gross financial wealth. "Private firm equity" denotes the book value of a household's holdings of private firm equity relative to lagged gross financial wealth. "Change in stock holdings" is the change in the value of the portfolio of directly held stocks given constant stock prices between $t-1$ and t relative to lagged gross financial wealth. "Change in h.h. debt" is the change in total household debt over lagged gross financial wealth. The first three outcomes in the bottom table are scaled by 100 for easier interpretation. Bootstrapped standard errors in parentheses. The rows labeled P(=Table 5) report the p-values from a bootstrapped t-test of equality between the coefficients in this table and in Table 5. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

Table A.10: Firm and Entrepreneur Outcomes – Sectoral Robustness

	Sales	Wage bill	Empl.	Value Added	VA/worker	EBITDA
$r_{i,t}$	82.4**	13.5+	0.05*	31.4*	−4.5	3.3
	(28.6)	(7.5)	(0.02)	(15.2)	(3.8)	(2.9)
N	553	553	553	553	553	553
Mean	2149.3	687.9	2.4	1247.8	658.1	167.5
Median	1383	555.8	1	909.8	495.3	98.2
P(=Table 5)	0.64	0.75	0.70	0.75	0.73	0.79

	Tot. assets	Fixed assets	Wk. Cap.	Equity	Tot. Liab.
$r_{i,t}$	27.7**	3.0*	5.0**	3.1*	23.3**
	(8.9)	(1.2)	(1.9)	(1.5)	(7.4)
N	553	553	553	553	553
Mean	785.3	133.3	118.3	177.7	573.1
Median	458	69.5	50.7	136.3	318.8
P(=Table 5)	0.65	0.80	0.61	0.96	0.69

	Private firm equity	Change in stock holdings	Change in h.h. debt	Log of pre-entry earnings
$r_{i,t}$	1.01*	−0.31**	3.17*	−0.01
	(0.46)	(0.11)	(1.36)	(0.01)
N	553	553	553	521
Mean	43.03	-2.11	11.17	13.08
Median	21.84	0	-16.97	13.12
P(=Table 5)	0.82	0.85	0.36	0.54

Notes: The table corresponds to Table 5 except that the following sectors are removed: Electricity, gas, steam, and air conditioning supply, Water supply, Mining and quarrying, Financial and insurance activities, Agriculture, forestry and fishing, and Real estate activities, Other service activities and Administrative and support service activities. All monetary values are in thousands of 2010 NOK. Outcomes are winsorized at the 5th and 95th percentiles. Working capital is defined as the difference between current assets and current liabilities. Entrepreneur balance sheet outcomes in the bottom panel are relative to lagged gross financial wealth. "Private firm equity" denotes the book value of a household's holdings of private firm equity relative to lagged gross financial wealth. "Change in stock holdings" is the change in the value of the portfolio of directly held stocks given constant stock prices between $t-1$ and t relative to lagged gross financial wealth. "Change in h.h. debt" is the change in total household debt over lagged gross financial wealth. The first three outcomes in the bottom table are scaled by 100 for easier interpretation. The results are based on the selection correction procedure described in Section 6. Bootstrapped standard errors in parentheses. The rows labeled P(=Table 5) report the p-values from a bootstrapped t-test of equality between the coefficients in this table and in Table 5. + denotes significance at the 10% level, * denotes significance at the 5% level, and ** denotes significance at the 1% level.

B Model Appendix

In this appendix, we present the details omitted from Section 2. First, we describe two example models both of which satisfy Assumption (M): one with financial frictions and one in which entrepreneurship provides non-pecuniary benefits. We also show that the two models differ in terms of their predictions for the effect of wealth on profits, which we estimate in our empirical analysis to differentiate the two models. Second, we show that our analysis is robust to allowing for ex-post residuals that might affect firm size and profits. Finally, we present the proofs omitted from the main text.

B.1 Models that Satisfy Assumption (M)

Consider the entry model described in the main text. Specifically, a continuum of individuals i differ in productivity z_i and initial assets a_i (along with observable covariates x_i). Conditional on entry, individuals' profits, capital, and non-pecuniary benefit from entrepreneurship are given by functions that depend only on their productivity and assets, $\pi_i = \pi\left(z_i, a_i\right), k_i = k\left(z_i, a_i\right), u_i^e = u^e\left(z_i, a_i\right)$. Individuals enter if their profits and non-pecuniary benefits from entrepreneurship exceed their reservation wage

$$\pi\left(z_i, a_i\right) + u^e\left(z_i, a_i\right) \geq w(z_i).$$

In the main text, we show that if $\pi\left(\cdot\right), k\left(\cdot\right), u^e\left(\cdot\right)$ and $w(\cdot)$ satisfy Assumption (M), which we reproduce here, then the model satisfies a rank preservation property that enables us to match entrants by productivity (without directly observing productivity).

Assumption (M). $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{da_i} \geq 0$ and $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i)-w(z_i))}{dz_i} > 0, \frac{dk(z_i,a_i)}{dz_i} > 0.$

We next describe two example economies that satisfy Assumption (M) but differ in terms of their predictions for $\frac{d\pi(z_i,a_i)}{da_i}$. In both cases, we assume the reservation wage does not depend on productivity, $w(z_i) = w$. We can allow productivity to increase wages to some extent, as long as the effect of productivity on wages is smaller than its effect on the net gain from entrepreneurship (which is strictly positive in both of our examples).

B.1.1 Model with Financial Frictions

Suppose there is no non-pecuniary benefit from entrepreneurship, $u^e\left(z_i, a_i\right) = 0$, but the entrepreneur might face financial frictions. In particular, entrepreneurs maximize their profits and they enter into business only if their maximum potential profit exceeds their reservation wage, $\pi\left(z_i, a_i\right) > w\left(z_i\right) = w$. We next describe the entrepreneur's production technology subject to financial frictions. We characterize the size and profit functions and show that they satisfy Assumption (M).

Suppose that if individual i starts a business, she produces according to the Cobb-Douglas technology

$$f(k; z_i) = z_i k^\alpha - rk - \kappa(z_i).$$

Here, r is the rental rate of capital and $\kappa(z_i)$ is a fixed entry cost. We assume entry costs are weakly decreasing in productivity $\kappa'(z_i) \leq 0$ (more productive individuals are able to reduce the fixed costs). For now, we assume capital is the only factor of production, which simplifies the algebra. The analysis can be extended to include labor.[1]

We capture financial constraints with a working capital channel. Specifically, the costs, $rk + \kappa(z_i)$, must be paid up front. The individual can use internal resources, a_i, to cover some of this cost. Therefore, the individual's borrowing need is $b = \max(0, rk + \kappa(z_i) - a_i)$. Individuals can obtain costly outside financing. Borrowing $b \geq 0$ costs $\phi_i(b)$ where $\phi_i(0) = 0$, $\phi_i'(b) > 0$ and $\phi_i''(b) > 0$. For simplicity, we work with the quadratic function, $\phi_i(b) = \phi(z_i)\frac{b^2}{2}$ where $\phi(z_i)$ is a constant. The quadratic functional form is not necessary for the qualitative results. Importantly, we assume the cost of financing, $\phi(z_i)$, is *weakly decreasing* in z_i.[2] More productive entrepreneurs obtain financing at a lower cost. This can be microfounded with a model in which there is default due to ex-post productivity shocks and outside financing costs depend on the likelihood of default (e.g., a costly-state verification model). In that type of model, a higher ex-ante productivity translates into a lower probability of default and therefore lower outside financing costs.[3]

With these assumptions, an entrepreneur that chooses to enter with capital k makes profits

$$\Pi(k; z_i, a_i) = z_i k^\alpha - rk - \kappa(z_i) - \phi(z_i)\frac{\max(0, rk + \kappa(z_i) - a_i)^2}{2}.$$

[1] If the labor bill is not subject to the working capital constraint (that we describe below), then labor is straightforward to incorporate. Specifically, suppose the production function is $\tilde{z}_i k^{\tilde{\alpha}} l^\theta - \kappa(z_i)$ and labor is supplied at a competitive wage w. Then, the entrepreneur always chooses the optimum amount of labor conditional on the other factors: that is

$$l^{opt} = \arg\max_l \tilde{z}_i k^{\tilde{\alpha}} l^\theta - \kappa(z_i) - rk - wl.$$

The solution is given by $l^{opt} = \left(\frac{\tilde{z}_i \theta k^{\tilde{\alpha}}}{w}\right)^{1/1-\theta}$. Substituting this back into the production function, we obtain

$$\begin{aligned} f(k; z_i) &= z_i k^\alpha - \kappa(z_i) - rk \\ \text{where } z_i &= \tilde{z}_i \left(\frac{\tilde{z}_i \theta}{w}\right)^{\theta/1-\theta} \text{ and } \alpha = \frac{\tilde{\alpha}}{1-\theta}. \end{aligned}$$

If labor is also subject to a working capital constraint, then incorporating labor would leave the results qualitatively unchanged but the algebra would be more complicated, since the firm would be optimizing over two factors.

[2] For the results of Lemma 1 we need to assume that either $\kappa(z_i)$ or $\phi(z_i)$ are strictly decreasing in z_i.

[3] We abstract away from asymmetric information on the ex-ante productivity z_i.

The entrepreneur's optimal profit and size (conditional on entry) are given by

$$\begin{aligned}\pi\left(z_i,a_i\right) &= \max_k \Pi\left(k;z_i,a_i\right)\\ k\left(z_i,a_i\right) &= \arg\max_k \Pi\left(k;z_i,a_i\right).\end{aligned}$$

The following result characterizes the comparative statics of the solution. The result also implies that this model satisfies Assumption (M) (recall that we assume the reservation wage is constant, $w(z_i)=w$).

Lemma 1. *Greater productivity strictly increases profits and size,* $\frac{d\pi(z_i,a_i)}{dz_i}>0, \frac{dk(z_i,a_i)}{dz_i}>0$. *In addition, greater initial assets weakly increase profits and size* $\frac{d\pi(z_i,a_i)}{da_i}\geq 0, \frac{dk(z_i,a_i)}{da_i}\geq 0$, *with strict inequality as long as the financial constraint binds.*

Proof of Lemma. First, consider the problem with $\phi\left(z_i\right)=0$: the first-best case without financial constraints. Denote the solution for this case with $\pi^*\left(z_i\right), k^*\left(z_i\right)$ (note that assets do not affect the solution in this case). Note that $\frac{d\pi^*(z_i)}{dz_i}>0, \frac{dk^*(z_i)}{dz_i}>0$.

Now consider the original problem. Consider the funding necessary to operate the business at the first-best level

$$\overline{a}\left(z_i\right)=rk^*\left(z_i\right)+\kappa\left(z_i\right).$$

There are two cases to consider. If $a_i>\overline{a}\left(z_i\right)$, the entrepreneur is effectively unconstrained and the problem is the same as the first-best case. If $a_i<\overline{a}\left(z_i\right)$, the entrepreneur is constrained. At an optimum point, the constraint binds and her profits are given by

$$\Pi\left(k;z_i,a_i\right)=z_ik^\alpha-rk-\kappa\left(z_i\right)-\frac{\phi\left(z_i\right)}{2}\left(rk+\kappa\left(z_i\right)-a_i\right)^2. \tag{B.1}$$

First consider the comparative statics of the optimal size, $k\left(z_i,a_i\right)$. The first order condition implies that $k\left(z_i,a_i\right)$ solves:

$$\frac{\partial\Pi\left(k;z_i,a_i\right)}{\partial k}=0\Longrightarrow z_i\alpha k^{\alpha-1}=r+\phi\left(z_i\right)\left(rk+\kappa\left(z_i\right)-a_i\right). \tag{B.2}$$

Implicitly differentiating with respect to a_i, we obtain

$$\frac{dk}{da_i}=-\frac{\frac{\partial^2\Pi(k;z_i,a_i)}{\partial k\partial a_i}}{\frac{\partial^2\Pi(k;z_i,a_i)}{\partial k^2}}>0. \tag{B.3}$$

Here, the inequality follows since $\frac{\partial^2\Pi(k;z_i,a_i)}{\partial k\partial a_i}=\phi\left(z_i\right)>0$ and Π is a concave function. Likewise, we have

$$\frac{dk}{dz_i}=-\frac{\frac{\partial^2\Pi(k;z_i,a_i)}{\partial k\partial z_i}}{\frac{\partial^2\Pi(k;z_i,a_i)}{\partial k^2}}>0, \tag{B.4}$$

since $\frac{\partial^2\Pi(k;z_i,a_i)}{\partial k\partial z_i} = -\phi'(z_i)(rk+\kappa-a_i) - \phi(z_i)\kappa'(z_i) > 0$.

Next consider the comparative statics of the optimal profit, $\pi(z_i,a_i)$. Using the Envelope Theorem, we obtain

$$\begin{aligned}\frac{d\pi(z_i,a_i)}{dz_i} &= \frac{\partial\Pi(k;z_i,a_i)}{\partial z_i}|_{k=k(z_i,a_i)}\\ &= k^\alpha - \frac{\phi'(z_i)}{2}(rk+\kappa-a_i)^2 > 0\\ \frac{d\pi(z_i,a_i)}{da_i} &= \frac{\partial\Pi(k;z_i,a_i)}{\partial a_i}|_{k=k(z_i,a_i)}\\ &= \phi(z_i)(rk+\kappa-a_i) > 0.\end{aligned}$$

Combining the two unconstrained and the constrained cases establishes the comparative statics and completes the proof. □

B.1.2 Model in which Entrepreneurship Provides Non-pecuniary Benefits

Consider the same model without financial frictions, $\phi(z_i) = 0$. Instead, suppose the non-pecuniary utility from entrepreneurship is given by a function of size and consumption

$$u^e = U^e(k,c;z_i,a_i) \text{ where } c = a_i + \Pi(k;z_i).$$

Here, c is consumption and $\Pi(k;z_i)$ denotes the profit function described in $(B.1)$ (we dropped the dependence on a_i since $\phi(z_i) = 0$). We assume the benefit from entrepreneurship satisfies

$$\frac{dU^e}{dc} > 0, \frac{dU^e}{dk} \geq 0, \frac{d^2U^e}{dcdk} \geq 0.$$

These assumptions capture the idea that individuals enjoy running a (larger) business, and more so when their regular consumption is higher. We also assume U^e is jointly concave in c and k and strictly concave in k. One example function that satisfies these assumptions is $U^e(k,c) = k^\gamma c^\beta$ for arbitrary $\gamma \in [0,1)$ and $\beta \in (0,1)$.

In this case, the entrepreneur solves

$$\begin{aligned}k(z_i,a_i) &= \arg\max_{k,c} c + U^e(k,c;z_i,a_i)\\ &\quad \text{s.t. } c = a_i + \Pi(k;z_i).\end{aligned} \tag{B.5}$$

The following lemma characterizes the solution and its comparative statics. The result also implies that this model satisfies Assumption (M) (recall that we assume the reservation wage is constant, $w(z_i) = w$).

Lemma 2. *Consider problem $(B.5)$ with the assumptions described above. The optimal size is the* *unique solution to*

$$\frac{d\Pi}{dk}\left(1+\frac{dU^e}{dc}\right)=-\frac{dU^e}{dk}. \tag{B.6}$$

The optimal size weakly exceeds the profit-maximizing size: that is, $k \geq k^*$ *where* $k^* = \arg\max_k \Pi(k, z_i)$. *Greater productivity increases the firm size and the total utility from entrepreneurship,* $\frac{dk(z_i,a_i)}{dz_i} > 0$, $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{dz_i} > 0$. *Greater wealth weakly increases the firm size and the total entrepreneurship utility,* $\frac{dk(z_i,a_i)}{da_i} \geq 0$, $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{da_i} \geq 0$, *but it* weakly decreases *firm profits,* $\frac{d\pi(z_i,a_i)}{da_i} \leq 0$.

Proof of Lemma. In view of the concavity assumptions, problem $(B.5)$ has a unique solution characterized by the optimality condition $(B.6)$. Since $\frac{dU^e}{dk} \geq 0$ and $\frac{dU^e}{dc} > 0$, this condition implies $\frac{d\Pi}{dk} \leq 0$. Since Π is strictly concave and the profit-maximizing size level k^* satisfies $\frac{d\Pi(k^*)}{dk} = 0$, this also implies $k \geq k^*$.

We next establish the comparative statics. Consider $\frac{dk(z_i,a_i)}{dz_i}$. We rewrite $(B.6)$ as $\frac{d(c+U^e)}{dk} = 0$. Implicitly differentiating this expression with respect to z_i, we obtain

$$\frac{dk(z_i,a_i)}{dz_i}=\frac{\frac{d^2(c+U^e)}{dkdz_i}}{-\frac{d^2(c+U^e)}{dk^2}}.$$

The denominator is strictly positive. Using $c = a_i + \Pi(k; z_i)$, we calculate the numerator as

$$\frac{d^2(c+U^e)}{dkdz_i}=\frac{d^2\Pi}{dkdz_i}\left(1+\frac{dU^e}{dc}\right)+\frac{d\Pi}{dk}\frac{d^2U^e}{dc^2}\frac{d\Pi}{dz_i}+\frac{d^2U^e}{dkdc}\frac{d\Pi}{dz_i}>0.$$

Here, the inequality follows since $\frac{d^2\Pi}{dkdz_i} > 0$, $\frac{dU^e}{dc} > 0$, $\frac{d^2U^e}{dc^2} \leq 0$, $\frac{d\Pi}{dk} \leq 0$, $\frac{d\Pi}{dz_i} > 0$ and $\frac{d^2U^e}{dkdc} \geq 0$. This proves $\frac{dk(z_i,a_i)}{dz_i} > 0$.

Next consider $\frac{dk(z_i,a_i)}{da_i}$. As before, $\frac{dk(z_i,a_i)}{da_i}$ has the same sign as $\frac{d^2(c+U^e)}{dkda_i}$. We calculate

$$\frac{d^2(c+U^e)}{dkda_i}=\frac{d\Pi}{dk}\frac{d^2U^e}{dc^2}+\frac{d^2U^e}{dkdc}\geq 0.$$

Here, the inequality follows since $\frac{d\Pi}{dk} \leq 0$, $\frac{d^2U^e}{dc^2} \leq 0$ and $\frac{d^2U^e}{dkdc} \geq 0$. This implies $\frac{dk(z_i,a_i)}{da_i} \geq 0$. Since $\frac{d\Pi}{da_i} = \frac{d\Pi}{dk}\frac{dk}{da_i}$ and $\frac{d\Pi}{dk} \leq 0$, this also implies $\frac{d\pi(z_i,a_i)}{da_i} \leq 0$.

Finally, consider the comparative statics of the total utility from entrepreneurship, $\pi(z_i, a_i) + u^e(z_i, a_i)$. Increasing z_i strictly increases the objective function in problem $(B.5)$ for any given choice of k. Therefore, it also strictly increases the maximum, which is given by $a_i + \Pi(k; z_i) + U^e(k; z_i)$. This implies $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{dz_i} > 0$. The same argument also implies $\frac{d(\pi(z_i,a_i)+u^e(z_i,a_i))}{da_i} \geq 0$, completing the proof. □

B.2 The model with unobserved residual heterogeneity

In the main text, we assumed firm size, profits, and outside options depend only on entrepreneurial productivity and assets. In practice, these variables might be heterogeneous also on other dimensions. In this appendix, we show that our selection approach is robust to allowing for residual heterogeneity under two conditions. First, we require the residual heterogeneity to be independent from initial wealth and entrepreneurial productivity conditional on observed characteristics. Second, we focus on entrants with size levels that exceed the entry cutoff by a margin. Under these assumptions, along with appropriate technical assumptions, we show that our rank matching approach still controls for average productivity and identifies the causal effect of initial wealth on firm profits and size.

Let $\overline{\delta}$ denote a vector of *ex-ante* characteristics that can influence firm size, profits, and outside options beyond entrepreneurial productivity z and initial wealth a. For instance, $\overline{\delta}$ can correspond to industry differences in technology or wage earnings potential that is orthogonal to entrepreneurial ability. Let $\overline{\varepsilon}$ denote a vector of *ex-post* characteristics that can influence ex-post firm size and profits beyond z and a. For instance, $\overline{\varepsilon}$ can correspond to ex-post productivity shocks or simply mistakes (relative to the optimal choice). We let $\tilde{k}\left(z,a,\overline{\delta},\overline{\varepsilon}\right),\tilde{\pi}\left(z,a,\overline{\delta},\overline{\varepsilon}\right),\tilde{u}^e\left(z,a,\overline{\delta},\overline{\varepsilon}\right)$ denote the size, profit, and entrepreneurial utility functions incorporating the ex-ante and ex-post residuals. We define the expected size, profit, and entrepreneurial utility functions as $k\left(z,a,\overline{\delta}\right)=E\left[\tilde{k}\left(z,a,\overline{\delta},\overline{\varepsilon}\right)\right],\pi\left(z,a,\overline{\delta}\right)=E\left[\tilde{\pi}\left(z,a,\overline{\delta},\overline{\varepsilon}\right)\right]$ and $\tilde{u}^e\left(z,a,\overline{\delta}\right)=E\left[\tilde{\pi}\left(z,a,\overline{\delta},\overline{\varepsilon}\right)\right]$. We let $w\left(z,\overline{\delta}\right)$ denote the outside option function that incorporates the ex-ante residuals.

Our main assumption is that $\overline{\delta}$ and $\overline{\varepsilon}$ are both independent from (z,a) and from each other *conditional* on x. Specifically, letting $F_{\overline{\delta},\overline{\varepsilon},z,a}$ denote the joint cumulative distribution of these variables, we strengthen Assumption (CIA) as follows.

Assumption (CIA$^{\mathrm{R}}$). As before, z and a are independent conditional on x. In addition, residual characteristics $\overline{\delta}$ and $\overline{\varepsilon}$ are independent from (z,a) and from each other conditional on x, that is: $F_{\overline{\delta}}\left(\overline{\delta}|\overline{\varepsilon},z,a,x\right)=F_{\overline{\delta}}\left(\overline{\delta}|x\right)$ and $F_{\overline{\varepsilon}}\left(\overline{\varepsilon}|\overline{\delta},z,a,x\right)=F_{\overline{\varepsilon}}\left(\overline{\varepsilon}|x\right)$.

For the rest of the appendix, we focus on individuals with a given x. At the expense of additional analytical complexity, we could combine Assumption (CIA$^{\mathrm{R}}$) and propensity score reweighting to generalize our results for the unconditional case, similar to how we extend Proposition 2 to Proposition 3 in the main text.

Focusing on individuals with a given x enables us to drop x from the notation. In particular, we define the *unobserved* residuals as the surprise component of the residuals given x:

$$\delta=\overline{\delta}-E\left[\overline{\delta}|x\right]\text{ and }\varepsilon=\overline{\varepsilon}-E\left[\overline{\varepsilon}|x\right].$$

Note that unobserved residuals have a zero mean by definition. We also redefine the functions

$\tilde{k}, \tilde{\pi}, k, \pi, w$ in terms of δ, ε as opposed to $\overline{\delta}, \overline{\varepsilon}$, for instance:

$$\tilde{k}(z, a, \delta, \varepsilon) = \tilde{k}\left(z, a, \delta + E\left[\overline{\delta}|x\right], \varepsilon + E\left[\overline{\varepsilon}|x\right]\right).$$

Note that these redefined functions implicitly depend on the covariates x. Likewise, we define the marginal distributions of unobserved residuals as

$$F_\delta(\delta) = F_{\overline{\delta}}\left(\delta + E\left[\overline{\delta}|x\right]|x\right) \text{ and } F_\varepsilon(\varepsilon) = F_{\overline{\varepsilon}}\left(\varepsilon + E\left[\overline{\varepsilon}|x\right]|x\right).$$

We assume the distributions $F_\delta(\delta), F_\varepsilon(\varepsilon)$ have bounded supports denoted by $\mathcal{D} = \left[-\overline{\delta}, \overline{\delta}\right]^{\mathcal{N}_\delta}$ and $\mathcal{E} = \left[-\overline{\varepsilon}, \overline{\varepsilon}\right]^{N_\varepsilon}$ (where N_δ and N_ε denote the size of the vectors δ and ε).

With residual heterogeneity, the individual chooses to enter into business, $E = 1$, if the following condition holds:

$$\pi(z, a, \delta) + u^e(z, a, \delta) \geq w(z, \delta).$$

We assume these functions and $k(z, a, \delta)$ satisfy the following version of assumption (M).

Assumption ($\mathbf{M}^{\mathrm{R}}$). $\frac{d(\pi(z,a,\delta)+u^e(z,a,\delta))}{da} \geq 0$ and $\frac{d(\pi(z,a,\delta)+u^e(z,a,\delta)-w(z,\delta))}{dz} > 0$, $\frac{dk(z,a,\delta)}{dz} > 0$.

With this assumption, a version of Proposition 1 still holds. Specifically, for each level of ex-ante residuals δ, there is a threshold level of productivity $\overline{z}(a, \delta)$ such that an individual enters into business if $z \geq \overline{z}(a, \delta)$. Since the total benefit from entry is increasing in initial assets a, we also have that $\overline{z}(a, \delta)$ is weakly decreasing in a. Therefore, for an initial wealth level a, the fraction of entrants is given by

$$e(a) = \int_{\mathcal{D}} \int_{z \geq \overline{z}(a,\delta)} dF_z(z)\, dF_\delta(\delta)$$

and it is weakly increasing in a, generalizing Proposition 1 to this case. Intuitively, residual heterogeneity that is uncorrelated with initial wealth does not change the prediction that higher initial wealth increases entry.

Our results on the effect of wealth on business characteristics require further adjustment. We can no longer use size to control for productivity, because size depends on both productivity and residuals. We next show how to adjust our procedure to account for residual heterogeneity.

A first issue is that unlike in the main text there is not a single productivity level corresponding to a given size level. Specifically, for a given k and a, this size might be chosen by entrepreneurs with lower z and higher ε, δ or higher z and lower ε, δ (assuming the convention that higher ε, δ increases k). To address this issue, we define the average productivity of the entrants that have initial wealth a and choose size k as:

$$z(k, a) = E\left[z|E = 1 \text{ and } \tilde{k}(z, a, \delta, \varepsilon) = k\right]. \tag{B.7}$$

Likewise, we also define the average profits for this group as

$$\pi(k,a) = E\left[\pi(z,a)\,|E=1 \text{ and } \tilde{k}(z,a,\delta,\varepsilon)=k\right].$$

A more major issue is that the distribution of size does not fully inform about the distribution of the (unobserved) productivity. To see this, observe that the fraction of entrants with size that exceeds k are now given by:

$$e\left(a,\tilde{k}\geq k\right) = \int_{\mathcal{D}}\int_{\mathcal{E}}\int_{E=1,\tilde{k}(z,a,\delta,\varepsilon)\geq \overline{k}} dF_z(z)\,dF_\varepsilon(\varepsilon)\,dF_\delta(\delta).$$

Unlike in the main text, this fraction is not necessarily equal to $\Pr(z\geq z(k,a)) = \int_{z\geq z(k,a)} dF_z(z)$, since it is also influenced by the distributions of δ and ε. Consequently, we cannot use the fractions $e\left(a,\tilde{k}\geq k\right)$ to match high and low initial wealth groups by productivity.

To make progress, we impose additional structure.

Assumption (L). For a given asset level a the size and the profit functions are linear in the remaining variables:

$$\tilde{k}(z,a,\delta,\varepsilon) = K(a) + K_z(a)\,z + K_\delta(a)\,\delta + K_\varepsilon(a)\,\varepsilon, \tag{B.8}$$

$$\tilde{\pi}(z,a,\delta,\varepsilon) = \Pi(a) + \Pi_z(a)\,z + \Pi_\delta(a)\,\delta + \Pi_\varepsilon(a)\,\varepsilon. \tag{B.9}$$

Here, $K(a), K_z(a), \Pi(a), \Pi_z(a)$ denote arbitrary constants and $K_\delta(a), \Pi_\delta(a), K_\varepsilon(a), \Pi_\varepsilon(a)$ denote conforming vectors. We assume $K_z(a), \Pi_z(a) > 0$ (consistent with Assumption (M)) and $K_\delta(a), K_\varepsilon(a)\geq 0$ (Convention).

Assumption (U). We also assume z has a uniform distribution: that is, $dF_z(\cdot)$ is constant.

These linearity and uniform-distribution assumptions enable us to obtain certainty-equivalence results and generalize our rank-matching approach, as long as we focus on size levels that exceed the entry cutoff by some margin.

Near the entry cutoffs, our approach might result in biased estimates due to selection driven by the entry decision. To see the issue, fix an initial wealth level a. Consider the cutoff productivity corresponding to this wealth level and the average residual, $\overline{z}(a,\mathbf{0})$. Let $\overline{k}(a) = K(a) + K_z(a)\,\overline{z}(a,\mathbf{0})$ denote the corresponding expected size. Consider the *entrants* with size $\overline{k}(a)$. The average productivity of these entrants is not necessarily equal to $\overline{z}(a,\mathbf{0})$, because some of the lower productivity individuals that would choose this size might not enter. Specifically, the lowest-productivity individual that chooses this size has high residuals $\overline{\delta},\overline{\varepsilon}$ and low productivity given by

$$\begin{aligned}\underline{z} &= \frac{\overline{k}(a) - K(a) - K_\delta(a)\,\overline{\delta} - K_\varepsilon(a)\,\overline{\varepsilon}}{K_z(a)}\\ &= \overline{z}(a,\mathbf{0}) - \frac{K_\delta(a)\,\overline{\delta} + K_\varepsilon(a)\,\overline{\varepsilon}}{K_z(a)}.\end{aligned}$$

This individual will enter only if her productivity satisfies:

$$\overline{z}(a,\mathbf{0}) - \frac{K_\delta(a)\overline{\delta} + K_\varepsilon(a)\overline{\varepsilon}}{K_z(a)} > \overline{z}\left(a,\overline{\delta}\right). \tag{B.10}$$

This condition is violated when there is only ex-post heterogeneity ($\overline{\delta} = 0$ and $\overline{\varepsilon} > 0$). The condition is also violated when there is ex-ante heterogeneity that affects size choice but not the entry decisions, $\overline{z}\left(a,\overline{\delta}\right) = \overline{z}(a,\mathbf{0})$. In these cases, and in many others, the average productivity of entrants with size $\overline{k}(a)$ will be *greater* than $\overline{z}(a,\mathbf{0})$. Intuitively, since size is driven by both residuals and productivity, and higher productivity agents are more likely to enter, the entrants with a given size will tend to be selected relatively more on productivity. Since this selection is unobserved and might be different for high and low-wealth groups, we cannot directly apply our procedure.

To circumvent this selection problem, we consider size levels that exceed the (average) entry cutoff by a sufficient margin so that all individuals that would choose this size enter. Formally, we assume:

$$\begin{aligned} k &> \overline{k}(a) + m(a) \qquad \text{(B.11)}\\ \text{where } \overline{k}(a) &= K(a) + K_z(a)\,\overline{z}(a,\mathbf{0})\\ \text{and } m(a) &= K_\varepsilon(a)\,\overline{\varepsilon} + K_\delta(a)\,\overline{\delta} + K_z(a)\left(\overline{z}\left(a,\overline{\delta}\right) - \overline{z}(a,\mathbf{0})\right). \end{aligned}$$

With this assumption, condition $(B.10)$ holds so the selection problem does not arise. Note also that the required margin $m(a)$ is decreasing in residual heterogeneity and limits to zero as $\overline{\varepsilon}, \overline{\delta} \to 0$. In the empirical implementation, we set $m(a)$ to a sizeable fraction of $\overline{k}(a)$. This is sufficient to address the selection problem as long as the residual heterogeneity is not too large.

Consider a size level $k > \overline{k}(a) + m(a)$. We next show that the average productivity of entrants with size k satisfies a certainty-equivalence property:

$$z(k,a) = E\left[z|E=1, \tilde{k}(z,a,\delta,\varepsilon) = k\right] = \frac{k - K(a)}{K_z(a)}. \tag{B.12}$$

In particular, $z(k,a)$ is equal to the productivity of the entrant with size k for which the residuals are equal to their means, $\delta, \varepsilon = 0$ (see $(B.8)$). To see this, first note that $z(k,a) = E\left[z|\tilde{k}(z,a,\delta,\varepsilon) = k\right]$: we can drop the conditioning on $E = 1$ since $k > \overline{k}(a) + m(a)$ $(B.7)$. Next substitute for $\tilde{k}(z,a,\delta,\varepsilon)$ from Assumption (L) and calculate the conditional expectation as:

$$z(k,a) = \frac{\int_\delta \int_\varepsilon z \Pr(\delta)\Pr(\varepsilon)\Pr\left(z = \frac{k - K(a) - K_\delta(a)\delta - K_\varepsilon(a)\varepsilon}{K_z(a)}\right) d\varepsilon d\delta}{\int_\delta \int_\varepsilon \Pr(\delta)\Pr(\varepsilon)\Pr\left(z = \frac{k - K(a) - K_\delta(a)\delta - K_\varepsilon(a)\varepsilon}{K_z(a)}\right) d\varepsilon d\delta}$$

$$
\begin{aligned}
&= \frac{\int_\delta \int_\varepsilon \frac{k-K(a)-K_\delta(a)\delta-K_\varepsilon(a)\varepsilon}{K_z(a)} dF_\delta(\delta)\, dF_\varepsilon(\varepsilon)}{\int_\varepsilon dF_\delta(\delta)\, dF_\varepsilon(\varepsilon)} \\
&= \frac{k-K(a)}{K_z(a)}.
\end{aligned}
$$

Here, the second line uses the uniform distribution assumption ($\Pr(z) = dF_z$ is constant) and the last line uses the normalizations, $E[\delta] = E[\varepsilon] = 0$. This establishes the certainty-equivalence property in (B.12).

We next show that the average profits for entrants with size k (that exceeds $\overline{k}+m$) satisfies a similar certainty-equivalence property:

$$
\pi(k,a) = \tilde{\pi}(z(k,a), a, \mathbf{0}, \mathbf{0}) = \Pi(a) + \Pi_z(a)\left(\frac{k-K(a)}{K_z(a)}\right). \tag{B.13}
$$

In particular, $\pi(k,a)$ is equal to the profits earned by an entrant for which the productivity is $z(k,a) = \frac{k-K(a)}{K_z(a)}$ and the residuals are equal to their means, $\delta, \varepsilon = 0$ (see (B.9)). To show this, we follow similar steps as above to calculate

$$
\begin{aligned}
\pi(k,a) &= \frac{\int_\delta \int_\varepsilon [\Pi(a) + \Pi_z(a) z + \Pi_\varepsilon(a)\varepsilon] \Pr(\delta)\Pr(\varepsilon)\Pr\left(z = \frac{k-K(a)-K_\varepsilon(a)\varepsilon}{K_z(a)}\right) d\varepsilon d\delta}{\int_\delta \int_\varepsilon \Pr(\delta)\Pr(\varepsilon)\Pr\left(z = \frac{k-K(a)-K_\delta(a)\delta-K_\varepsilon(a)\varepsilon}{K_z(a)}\right) d\varepsilon d\delta} \\
&= \int_\delta \int_\varepsilon \left(\Pi(a) + \Pi_z(a)\left(\frac{k-K(a)-K_\delta(a)\delta-K_\varepsilon(a)\varepsilon}{K_z(a)}\right) + \Pi_\varepsilon(a)\varepsilon\right) dF_\delta(\delta)\, dF_\varepsilon(\varepsilon) \\
&= \Pi(a) + \Pi_z(a)\left(\frac{k-K(a)}{K_z(a)}\right).
\end{aligned}
$$

This establishes the certainty-equivalence property in (B.13).

Finally, we show that the fraction of entrants with size that exceeds k satisfies a similar certainty-equivalence property:

$$
e\left(a, \tilde{k} \geq k\right) = \int_{z \geq z(k,a)} dF_z(z). \tag{B.14}
$$

In particular, the fraction $e\left(a, \tilde{k} \geq k\right)$ is equal to the fraction that would obtain in a model in which the residuals are constant and equal to their means, $\delta, \varepsilon = 0$. To see this, note that:

$$
\begin{aligned}
e\left(a, \tilde{k} \geq k\right) &= \int_\varepsilon \int_{K(a)+K_z(a)z+K_\varepsilon(a)\varepsilon \geq k} dF_z(z)\, dF_\delta(\delta)\, dF_\varepsilon(\varepsilon) \\
&= \int_\varepsilon \Pr\left(z \geq \frac{k-K(a)-K_\delta(a)\delta-K_\varepsilon(a)\varepsilon}{K_z(a)}\right) dG(\varepsilon) \\
&= \Pr\left(z \geq \frac{k-K(a)}{K_z(a)}\right) = \int_{z \geq z(k,a)} dF_z(z).
\end{aligned}
$$

Here, the last line uses the observation that the inverse cumulative distribution function $\Pr\left(z \geq \frac{k-K(a)-K_\delta(a)\delta-K_\varepsilon(a)\varepsilon}{K_z(a)}\right)$ is a linear function of its argument (since z is uniform) along with

$E[\varepsilon] = 0$. This establishes the certainty-equivalence property in $(B.14)$.

Eq. $(B.14)$ shows that with linearity and uniform-distribution assumptions we *can* still use the fractions $e\left(a, \tilde{k} \geq k\right)$ to control for the (unobserved) average productivity, $z(k,a)$. Thus, we have the following result that generalizes Proposition 2 to the case with residual heterogeneity.

Proposition 4 (Rank preservation with residual heterogeneity)**.** *Consider the model with (ex-ante or ex-post) residual heterogeneity. Suppose Assumptions (M^R), (L), and (U) hold. Consider an initial wealth level a^L and a corresponding size level k^L. Let $z^L = z\left(k^L, a^L\right)$ denote the average productivity of entrants with k^L and a^L. Let $a^H > a^L$ denote a higher wealth level. Suppose k^L is sufficiently high that the following inequalities both hold*

$$k^L > \overline{k}\left(a^L\right) + m\left(a^L\right) \text{ and } \tilde{k}\left(z^L, a^H, \mathbf{0}, \mathbf{0}\right) > \overline{k}\left(a^H\right) + m\left(a^H\right), \tag{B.15}$$

where the functions $\overline{k}(\cdot)$ and $m(\cdot)$ are given by $(B.11)$. Let $\overline{k} \geq k^L$ denote the unique size level that solves,

$$e\left(a^H, \tilde{k} \geq \overline{k}\right) = e\left(a^L, \tilde{k} \geq k^L\right). \tag{B.16}$$

Then, $\overline{k} = \tilde{k}\left(z^L, a^H, \mathbf{0}, \mathbf{0}\right)$: that is, $\overline{k}$ is the firm size an entrant with productivity z^L and average residuals $\delta, \varepsilon = 0$ would have if she had higher initial wealth (and the same productivity). Thus, comparing $\overline{k}$ and k^L identifies the causal effect of initial wealth on firm size:

$$\overline{k} - k^L = \tilde{k}\left(z^L, a^H, \mathbf{0}, \mathbf{0}\right) - \tilde{k}\left(z^L, a^L, \mathbf{0}, \mathbf{0}\right). \tag{B.17}$$

Likewise, comparing the average profits of the high-wealth entrants with size $\overline{k}$ with the average profits of the low-wealth entrants with size k^L identifies the causal effect of initial wealth on firm profits for an entrant with productivity z^L and average residuals $\delta, \varepsilon = 0$:

$$\pi\left(\overline{k}, a^H\right) - \pi\left(k^L, a^L\right) = \tilde{\pi}\left(z^L, a^H, \mathbf{0}, \mathbf{0}\right) - \tilde{\pi}\left(z^L, a^L, \mathbf{0}, \mathbf{0}\right). \tag{B.18}$$

Despite residual heterogeneity, our rank-matching approach controls for average productivity and identifies the causal effects for an entrant with average productivity and average residuals. The intuition follows from the certainty-equivalence properties we have established. While the residuals shuffle the firm size (for a given wealth), they do not bias our approach in a particular direction. Some larger firms are less productive, and some smaller firms are more productive, but the firm with the average size has the same average productivity as in the baseline case with certainty. Thus, matching firms by their size rank across high and low-wealth group still controls for average productivity.

While assumptions (L) and (U) enable us to obtain exact certainty-equivalence, this intuition suggests that the result is likely to hold approximately under weaker assumptions. In numerical simulations, we verify that our rank-matching approach controls for average productivity well also when z shocks follow a different distribution (e.g., a normal distribution) than uniform. In this sense, we view (L) and (U) as technical assumptions.

Proof of Proposition 4. Let $k^H = \tilde{k}\left(z^L, a^H, \mathbf{0}, \mathbf{0}\right) = K\left(a^H\right) + K_z\left(a^H\right) z^L$ denote the size for an entrant with initial wealth a^H, productivity z^L, and average residuals. Note that the inequalities in $(B.15)$ imply that assumption $(B.11)$ holds for both pairs $a = a^L, k = k^L$ and $a = a^H, k = k^H$. Note also that Eq. $(B.12)$ implies $z\left(k^H, a^H\right) = z^L$: the average productivity of entrants with initial assets a^H and size $k^H = K\left(a^H\right) + K_z\left(a^H\right) z^L$ is equal to z^L.

Then, applying Eq. $(B.14)$ both pairs $a = a^L, k = k^L$ and $a = a^H, k = k^H$ we obtain

$$\begin{aligned} e\left(a^L, \tilde{k} \geq k^L\right) &= \int_{z \geq z(k^L, a^L)} dF_z(z) = \int_{z \geq z^L} dF_z(z) \\ e\left(a^H, \tilde{k} \geq k^H\right) &= \int_{z \geq z(k^H, a^H)} dF_z(z) = \int_{z \geq z^L} dF_z(z). \end{aligned}$$

These expressions imply that k^H is the unique solution to $(B.16)$: that is, $\overline{k} = \tilde{k}\left(z^L, a^H, \mathbf{0}, \mathbf{0}\right)$. This in turn implies $(B.17)$.

Likewise, applying Eq. $(B.13)$ for both pairs $a = a^L, k = k^L$ and $a = a^H, k = k^H = \overline{k}$ we obtain

$$\begin{aligned} \pi\left(k^L, a^L\right) &= \tilde{\pi}\left(z\left(k^L, a^L\right), a^L, \mathbf{0}, \mathbf{0}\right) = \tilde{\pi}\left(z^L, a^L, \mathbf{0}, \mathbf{0}\right) \\ \pi\left(\overline{k}, a^H\right) &= \tilde{\pi}\left(z\left(\overline{k}, a^H\right), a^H, \mathbf{0}, \mathbf{0}\right) = \tilde{\pi}\left(z^H, a^L, \mathbf{0}, \mathbf{0}\right). \end{aligned}$$

Combining these expressions implies $(B.18)$, completing the proof. □

B.3 Omitted Proofs in Section 2

Proof of Proposition 1. Under Assumption (M), the net gain from entry, $\pi(z_i, a_i) + u^e(z_i, a_i) - w(z_i)$, is weakly increasing in a_i and strictly increasing in z_i. The latter relation implies that for any a_i there exists a threshold level $\overline{z}(a_i)$ such that an agent enters if and only if $z_i \geq \overline{z}(a_i)$. The former relation implies that $\overline{z}(a_i)$ is weakly decreasing in a_i, completing the proof. □

Proof of Proposition 2. We first claim that $\overline{k} = k\left(z, a^H\right)$ is the unique solution to Eq. (1). To this end, note that

$$\begin{aligned} e\left(a^L, k \geq k^L | x_i\right) &= \int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k^L} dF_z(z_i | x_i) \\ &= \int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k(z, a^L)} dF_z(z_i | x_i) \\ &= \int_{z_i \geq z} dF_z(z_i | x_i). \end{aligned}$$

Here, the second line substitutes $k_L = k\left(z, a^L\right)$ and the last line follows since $k\left(z_i, a^L\right)$ is monotonic in z_i (and $z_i \geq z$ implies $z_i \geq \overline{z}\left(a^L\right)$). The same steps imply that for $\overline{k} = k^H\left(z, a^H\right)$ we have,

$$\begin{aligned} e\left(a^H, k \geq \overline{k}|x_i\right) &= \int_{z_i \geq \overline{z}(a^H), k(z_i, a^H) \geq \overline{k}} dF_z\left(z_i|x_i\right) \\ &= \int_{z_i \geq \overline{z}(a^H), k(z_i, a^H) \geq k^H(z, a^H)} dF_z\left(z_i|x_i\right) \\ &= \int_{z_i \geq z} dF_z\left(z_i|x_i\right). \end{aligned}$$

Comparing these expressions proves that $\overline{k} = k^H\left(z, a^H\right)$ solves Eq. (1). Note also that $e\left(a^H, k \geq \overline{k}|x_i\right)$ is strictly decreasing in $\overline{k}$, because the function $k\left(z_i, a^H\right)$ is strictly increasing in z_i and the distribution $dF_z\left(z_i|x_i\right)$ is continuous in z_i. This implies that $\overline{k} = k\left(z, a^H\right)$ is the unique solution to Eq. (1).

Next consider the firm-level outcome $\overline{y}$ corresponding to the firm with size $\overline{k}$ and initial assets a^H. Since $\overline{k} = k\left(z, a^H\right)$, we also have $\overline{y} = y\left(z, a^H\right)$. This implies $\overline{y} - y^L = y\left(z, a^H\right) - y\left(z, a^L\right)$ and completes the proof. □

Proof of Proposition 3. We first claim that $\overline{k} = k\left(z, a^H\right)$ is the unique solution to (3). To this end, observe that

$$\begin{aligned} e^*\left(a^L, k \geq k^L\right) &= \int_{x_i}\int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k^L} dF_z\left(z_i|x_i\right)\omega\left(x_i\right) dF_x\left(x_i|a^L\right) \\ &= \int_{x_i}\int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k^L} dF_z\left(z_i|x_i\right) dF_x\left(x_i|a^H\right) \\ &= \int_{x_i}\int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k^L} f_z\left(z_i|x_i\right) f_x\left(x_i|a^H\right) dz_i dx_i \\ &= \int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k^L} \left(\int_{x_i} f_z\left(z_i|x_i\right) f_x\left(x_i|a^H\right) dx_i\right) dz_i \\ &= \int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k^L} f_z\left(z_i|a^H\right) dz_i \\ &= \int_{z_i \geq \overline{z}(a^L), k(z_i, a^L) \geq k(z, a^L)} f_z\left(z_i|a^H\right) dz_i \\ &= \int_{z_i \geq z} f_z\left(z_i|a^H\right) dz_i \end{aligned}$$

Here, the second line uses the definition of the propensity score $\omega\left(x_i\right) = \frac{dF_x\left(x_i|a^H\right)}{dF_x\left(x_i|a^L\right)}$, the third line substitutes the PDFs corresponding to the CDFs, the fourth line changes the order of integration, the fifth line substitutes the definition of the marginal PDF $f_z\left(z_i|a^H\right) = \int_{x_i} f_z\left(z_i|x_i\right) f_x\left(x_i|a^H\right) dx_i$, the sixth line substitutes $k^L = k\left(z_i, a^L\right)$, and the last line follows since $k\left(z_i, a^L\right)$ is monotonic in

z_i (and $z_i \geq z$ implies $z_i \geq \overline{z}\left(a^L\right)$). Following similar steps, for $\overline{k} = k\left(z, a^H\right)$, we have

$$
\begin{aligned}
e\left(a^H, k \geq \overline{k}\right) &= \int_{x_i} \int_{z_i \geq \overline{z}(a^H), k(z_i, a^H) \geq \overline{k}} dF_z\left(z_i | x_i\right) dF_x\left(x_i | a^H\right) \\
&= \int_{x_i} \int_{z_i \geq \overline{z}(a^H), k(z_i, a^H) \geq \overline{k}} f_z\left(z_i | x_i\right) f_x\left(x_i | a^H\right) dz_i dx_i \\
&= \int_{z_i \geq \overline{z}(a^H), k(z_i, a^H) \geq \overline{k}} \int_{x_i} f_z\left(z_i | x_i\right) f_x\left(x_i | a^H\right) dx_i dz_i \\
&= \int_{z_i \geq \overline{z}(a^H), k(z_i, a^H) \geq \overline{k}} f_z\left(z_i | a^H\right) dz_i \\
&= \int_{z_i \geq \overline{z}(a^H), k(z_i, a^H) \geq k(z, a^H)} f_z\left(z_i | a^H\right) dz_i \\
&= \int_{z_i \geq z} f_z\left(z_i | a^H\right) dz_i.
\end{aligned}
$$

Comparing these expressions proves that $\overline{k} = k\left(z, a^H\right)$ solves (3). Note also that $e\left(a^H, k \geq \overline{k}\right)$ is strictly decreasing in $\overline{k}$, which implies that $\overline{k} = k^H\left(z, a^H\right)$ is the unique solution to Eq. (3) .

Next consider the firm-level outcome $\overline{y}$ corresponding to the firm with size $\overline{k}$ and initial assets a^H. Since $\overline{k} = k\left(z, a^H\right)$, we also have $\overline{y} = y\left(z, a^H\right)$. This implies $\overline{y} - y^L = y\left(z, a^H\right) - y\left(z, a^L\right)$ and completes the proof. □