FINNISH CENTRE OF EXCELLENCE IN TAX SYSTEMS RESEARCH

FIT Working Paper 17

Jarkko Harju, Toni Juuti and Tuomas Matikka

Stairway to Heaven? Selection into Entrepreneurship, Income Mobility and Firm Performance

Tampere University
UNIVERSITY OF HELSINKI
VATT INSTITUTE FOR ECONOMIC RESEARCH
CENTRES OF EXCELLENCE IN RESEARCH

Stairway to Heaven? Selection into Entrepreneurship, Income Mobility and Firm Performance*

Jarkko Harju† Toni Juuti‡ Tuomas Matikka§

December 20, 2023

Abstract

Using detailed full-population data from Finland, we provide evidence on selection into entrepreneurship and the dynamic implications of establishing a new business. Individuals at the very top of the personal income distribution are much more likely to start a new incorporated business compared to others. There is no similar selection based on parental income, but more than half of new entrepreneurs have entrepreneurial parents. Entrepreneurship is associated with a similar average income gain of 20% relative to comparable wage earners throughout both personal and parental income distributions. However, key firm-level outcomes such as productivity and job creation are positively linked with personal income. This suggests that high-income individuals do not only benefit from entrepreneurship personally, but their businesses are associated with the largest positive spillovers in the economy. In contrast, we find no significant differences in the outcomes of new firms by parental income or parental background in entrepreneurship.

Keywords: entrepreneurship; income mobility; inequality; productivity.

JEL Codes: L26; J24; J3

*We thank Adrian Adermon, Jonas Hjort, Niels Johannesen, Markus Jäntti, Martti Kaila, Ross Levine, Matthew Lindquist, Oskar Nordström Skans, Martin Nybom, Tuuli Paukkeri, Tuomas Pekkarinen, Jukka Pirttilä, Emmanuel Saez, Matteo Tranchero, Joacim Tåg, Danny Yagan, Eric Zwick and several conference participants and discussants for their comments and suggestions. All remaining errors are our own. We are grateful to the Academy of Finland (grant no. 346252 and 340198) and Palkansaaja Foundation (Innovation, Productivity and Growth) for the funding for this project.

†Tampere University & VATT Institute for Economic Research & CESifo.

‡Labour Institute for Economic Research & Tampere University.

§VATT & CESifo.

1 Introduction

Entrepreneurs are pivotal for generating growth, innovation and employment opportunities in modern economies (see e.g. Aghion et al. 2018a; Decker et al. 2014). Entrepreneurs are often rewarded for their risk-taking, as business owners tend to occupy the top of the income distribution in many developed countries (see e.g. Smith et al. 2022, 2019; Halvarsson et al. 2018). However, despite a large literature on selection into entrepreneurship and differences in incomes between entrepreneurs and salaried workers (see e.g. Hamilton 2000; Moskowitz and Vissing-Jørgensen 2002; Hurst and Lusardi 2004; Levine and Rubinstein 2017; Kerr et al. 2017), we do not exactly know who benefits from entrepreneurship. Does becoming an entrepreneur increase the incomes of those who are already well-off, or does it provide a "Stairway to Heaven" for those who come from less privileged backgrounds?

This article provides comprehensive empirical evidence on selection into entrepreneurship and the dynamic implications of new business creation using population-wide administrative data linking together wage earners, business owners and their firms. Our study integrates various separate strands of literature on selection and income development among entrepreneurs (Kerr et al. 2017; Levine and Rubinstein 2017; Garnadt et al. 2023, among many others), the intergenerational mobility of business owners (Lindquist and Vladasel 2022), the role of individual (Hvide and Møen 2010; Queiro 2022) and parental (Hurst and Lusardi 2004; Hurst and Pugsley 2011; Lindquist et al. 2015) characteristics in entrepreneurial activity, and the role of entrepreneurship in income inequality (Kopczuk and Zwick 2020; Smith et al. 2019, 2022).

Our first contribution is to exploit unique and newly compiled data to form a precise timeline for transitions into entrepreneurship. Using these data, we provide a detailed and visually clear analysis of selection into entrepreneurship, and how becoming an entrepreneur shapes the income trajectories of individuals by following them long before and after they established their first business. We study these outcomes without conditioning on the survival or success of the business, in contrast to what has typically been done in previous studies.

Despite the extensive earlier literature on entrepreneurship, we do not yet know whether entrepreneurs dominate the top of the income distribution because of the positive associated income gains from entrepreneurship, or because high income individuals are typically more likely to become business owners in the first place.

Our second novelty is to link business owners to their newly established firms. We study the productivity and job creation of the new firms and reveal how individual and parental characteristics are associated with firm performance over time. To our knowledge, this paper is the first go beyond individual-level outcomes by studying both the income trajectories of new entrepreneurs and the success of new businesses by individual and parental background.

A necessary requirement for our study is the opportunity to combine detailed population-wide panel data sets from many sources. We link together administrative registers of wage earners, business owners, and their firms using data from Finland. These data contain information on key individual and firm-level outcomes, such as detailed income information on individuals and their parents, and the financial statements of the firms. We can use these data to define the start dates of the businesses and the ownership structures of firms, and to distinguish newly established businesses from existing ones. Importantly, we can construct a tractable and visually clear timeline around the event of founding a new business, enabling us to follow both individual and firm-level outcomes over a long time period before and after the firm was established.

Following recent literature, Levine and Rubinstein (2017) most notably, we split business owners into owners of incorporated and unincorporated firms, and primarily focus on incorporated entrepreneurs. Incorporated businesses constitute separate legal entities with limited liability, which encourage growth and risk-taking compared to the unlimited liability of the owners of unincorporated businesses. Earlier studies document significant heterogeneity in typical tasks, incomes and personality traits between the owners of different company forms, and typically find that owners of incorporated businesses end up earning much more than other business owners and wage earners (see e.g. Hurst and Pugsley 2011; Levine and

Rubinstein 2017, 2020; Lindquist and Vladasel 2022).

Our approach departs from the earlier related literature in one key aspect. Instead of asking how being an owner of a business is associated with personal income development, we ask how the decision to become an entrepreneur and establish a new business with no previous experience in running one shapes the income trajectories of individuals. Thus, we focus on new business creation instead of, for example, individuals joining existing companies as new owners. Also, our analysis is not conditional on firm success or "making it" as an entrepreneur, as we include individuals who choose to set up a new business but whose firms did not succeed. Most earlier studies rely on either cross-sectional data or survey-based panels with relatively few individuals followed over time, which tends to over-represent business owners who manage mature firms (see e.g. Hamilton 2000, Moskowitz and Vissing-Jørgensen 2002, Levine and Rubinstein 2017, Halvarsson et al. 2018). One notable exception is a study by Berglann et al. (2011), who follow one cohort of new entrepreneurs over time using Norwegian panel data. We provide novel results highlighting that focusing on new entrepreneurs instead of the entire pool of existing entrepreneurs, including those who join existing businesses as new owners, leads to very different implications on selection and income dynamics.

As our first result, we find that a large share of incorporated entrepreneurs were top earners already before they became entrepreneurs. More than 5% of new entrepreneurs came from the top percentile of the income distribution, while the top 5% included 17% of them. Previous literature has acknowledged a general positive connection between incomes and the likelihood of being an incorporated business owner (Levine and Rubinstein 2017, 2020; Kerr et al. 2017). Our more granular analysis highlights that selection into entrepreneurship occurs very non-linearly and only at the very top of the income distribution. Hurst and Lusardi (2004) document a similar pattern regarding personal wealth and entrepreneurship using survey data from the US.

In contrast to individual incomes, the likelihood of starting a new business decreases

slightly by parental income ranks. This illustrates that the association between parental income and becoming a new entrepreneur differs significantly from earlier findings (Dunn and Holtz-Eakin 2000; Levine and Rubinstein 2017; Lindquist and Vladasel 2022), which do not focus on *new* entrepreneurs but rather study the links between parental income and entrepreneurship including all existing business owners. Our results on this type of general correlation between all types of business ownership and parental income are similar to these earlier studies, finding a positive link between parental income and entrepreneurship. In contrast, our detailed analysis shows that this finding does not come from a positive connection between parental incomes and new business creation, but rather stems from the fact that children from high-income families are more likely to join an existing business as a new owner, and that their firms are more likely to survive in the longer run.

Selection into entrepreneurship by high-income individuals is commonly rationalized by entrepreneurial skills, liquidity or preferences for entrepreneurship (see e.g. Levine and Rubinstein 2020). We find that individual income is linked to the initial capital installments of new businesses, which increase almost linearly by income rank. However, as discussed above, selection by individual income occurs only at the very top of the distribution. Therefore, the liquidity channel cannot fully rationalize the observed selection pattern, leaving room for the interpretation that high-income individuals are positively selected by preferences or entrepreneurial ability, which is also well rewarded in salaried employment before establishing a business.

In contrast, we observe no association between parental income and the initial capital assets of new firms. This further suggests that parental income has no significant connection with new business creation. However, as in many previous studies (see e.g. Dunn and Holtz-Eakin 2000; Uusitalo 2001; Lindquist et al. 2015), parental background in entrepreneurship is heavily associated with starting a business. Nevertheless, we find that parental entrepreneurship does not drive the link between individual incomes and selection into starting a firm: the likelihood of becoming a new entrepreneur is pronounced at the very top of the earnings

distribution for both those with and without entrepreneurial parents.

We then analyze how entrepreneurship affects income mobility. The income gains from incorporated entrepreneurship compared to paid employment are on average 20% ten years after the new business was started. These gains are rather evenly distributed across both individual and parental income distributions. This means that even though top earners are much more likely to become entrepreneurs, the relative gains from entrepreneurship are not positively correlated with initial personal income rank. However, the gains are somewhat larger for those who have both high individual income and come from high-income families. These findings suggest that entrepreneurship is linked with upward income mobility across both individual and parental income distributions, as those who come from the bottom and middle parts of the distribution also gain significantly from entrepreneurship. However, we underline that these results are descriptive in nature and do not reflect the causal effect of entrepreneurship on personal income development. This is highlighted by the clear selection into entrepreneurship at the the top of the income distribution.

Combining our findings on selection and mobility demonstrates that entrepreneurs appear to occupy the top of the distribution even before starting a business due to positive self-selection by income, but their preponderance at the top is further enhanced by the positive income gains associated with founding an incorporated business. As many as 17% of new incorporated entrepreneurs were among the top 5% of all income earners in Finland already in the year before the firm was established, and 10 years after the firm was established this share increases to 26%. In the top 1% the share of entrepreneurs more than doubles from 5% to 11%.

To further illustrate the role of entrepreneurs at the top of the distribution, we find that business income and other capital income clearly dominate at the very top, similarly as in the US (Smith et al. 2019). When zooming into the top 0.1% of the market income distribution, the share of capital and business income is above 60% and wage income falls below 40% of aggregate market income. When we allocate the earnings retained in firms to

the incorporated business owners, the share of business and capital income of total market income further increases to 80% in the top 0.1%.

We extend our analysis from individual-level income development and mobility to firm-level outcomes. This provides a more complete view of the implications of new entrepreneurship, enabling us to capture how individual and parental characteristics are linked to the success of businesses and the potential positive spillovers of new business creation. Contrary to the evenly distributed gains across the personal income distribution, we find that measures of firms' success, such as value added and job creation, clearly increase with initial personal income. These findings imply that the positive externalities provided by high-income individuals are large relative to their personal gains. As we observed that the initial capital stock invested in new firm increases with individual income, at least part of the observed firm-level development by income ranks can be driven by initial assets instead of entrepreneurial ability. This suggests that liquidity and assets are important for the success of businesses, as emphasized in many earlier studies (see e.g. Evans and Leighton 1989; Evans and Jovanovic 1989; Holtz-Eakin et al. 1994; Fairlie 1999; Hurst and Lusardi 2004; Hvide and Møen 2010; Levine and Rubinstein 2020).

Earlier literature underlines the role of parental background and the associated human capital channel in explaining entrepreneurial traits and activity (see e.g. Dunn and Holtz-Eakin 2000; Hurst and Lusardi 2004; Hurst and Pugsley 2011; Lindquist et al. 2015). However, we find that none of the key business activity measures increase with parental income, including the initial capital stock of new firms. Furthermore, we find that firm-level outcomes are not linked with parental background in entrepreneurship either, as they develop very similarly for both those with and without entrepreneurial parents. Therefore, our findings suggest that despite the fact that entrepreneurship often runs in the family, second-generation entrepreneurs who establish new businesses do not appear to succeed better than those whose parents were not business owners. These findings, therefore, do not lend support to the popular idea that inheritable entrepreneurial traits or "inherited" human capital are

positively linked with entrepreneurial success.

Finally, using similar data restrictions as the previous literature, we find that many key observations, such as the key role of parental background in selection into entrepreneurship (Fairlie and Robb 2007; Lindquist et al. 2015), incorporated entrepreneurs' higher incomes (Levine and Rubinstein 2017; Halvarsson et al. 2018) and the high share of business income at the very top of the distribution (Smith et al. 2019; Bruil et al. 2022), are strikingly similar in Finland compared to the evidence from other countries, which strengthens the external validity of our results on selection into entrepreneurship, income mobility and firm performance.

The paper proceeds as follows: Section 2 describes the data and explains the detailed definitions and the empirical approach we use in our analysis. Section 3 presents the results on selection into entrepreneurship, and Section 4 the results on income dynamics and mobility. In Section 5, we characterize how selection and income dynamics contribute to top-end inequality. Section 6 presents the analysis on firm performance. Section 7 concludes.

2 Data and Definitions

2.1 Data

Our analysis builds on population-wide administrative data from Statistics Finland and the Finnish Tax Administration. These data enable us to link together individual-level income information and background characteristics, firm ownership information and date of establishment, and firm-level tax returns and financial statements. The data are unique both in terms of details and coverage.[1]

[1]We provide a brief comparison with data used in previous related studies in Appendix E. Previous studies on business owners using Finnish data have documented that entrepreneurship appears to be a character trait that runs in the family (Uusitalo 2001), that entrepreneurs' education is positively associated with entrepreneurial success and that smaller firms spawn new entrepreneurs more frequently than larger firms (Hyytinen and Maliranta 2008). Within-twin analysis (Hyytinen et al. 2013; Maczulskij and Viinikainen 2023) suggests that entrepreneurs earn a negative earnings premium, work more, have greater responsibilities, but also have greater control over their work, and that self-confidence predicts entry into entrepreneurship and

Individual-level Information. The individual-level data from Statistics Finland include detailed annual income information (wages, dividends, capital gains and other sources of earned and capital income), education (level and field of highest degree), age, place of residence, marital status, and the number of children. The data contain the universe of individuals living in Finland from 1987 to 2019 with unique individual-year identifiers. We can also link children to their parents (biological and adoptive) for all individuals born after 1952 and permanently living in Finland on the last day of each year.

Firms. Our firm-level data set includes annual financial statements (profit and loss accounts, balance sheet information) from Statistics Finland, firm-level tax return information from the Finnish Tax Administration, and the establishment dates of firms from the Finnish Patent and Registration Office from 1997 to 2019. The data include firm-year identifiers and additional background characteristics for firms such as industry classification codes and the organizational form.

Ownership Information. For the owners of privately held corporations, we use two separate data sets to identify the ownership structures: 1) the main owner information from the Tax Administration for those owners who received dividend income from their firm, available for 1997–2016, and 2) the full ownership database from Statistics Finland, available for 2006–2019. Combining these data enables us to track down the main owners with dividend income from their firm from 1997–2005 and the full ownership structure from 2006–2019.[2] Second, we use a data set including all self-employed and partnership firm owners from the

entrepreneurial success. Previous experience in a high-productivity firm strongly predicts high productivity and probability of survival for the entrepreneur's new firm (Maliranta and Nurmi 2019). Paukkeri et al. (2023) show that entrepreneurs dominate the top of the distribution in Finland, which is similar to our findings. Finally, recent studies on Finnish inventors (Aghion et al. 2018b, 2022, 2023) also use similar data sources as we do.

[2]Figure A1 in Appendix A shows the shares of new incorporated and unincorporated entrepreneurs by year. As is visible from the figure, for incorporated entrepreneurs there is a data break in 2006, as we only observe the main owners receiving dividends from their firm before 2006, and from then onwards we observe all main owners of incorporated firms. We use all the firm starts in our baseline analysis. Our main findings are not sensitive to using data before or after 2006, as shown Table B1 in Appendix B.

Tax Administration available for 1997–2016. Both of these data sets include unique firm-owner-year identifiers, and thus we can merge these data with all the individual-level data sets.

Sample Restrictions. Similarly as in the recent literature on the income development and personality traits of entrepreneurs (Levine and Rubinstein 2017; Halvarsson et al. 2018), we restrict our baseline sample to individuals who are 25–64 years old when establishing their business and to individuals with positive wage income and/or entrepreneurial income in a given year. We also restrict our main sample to those individuals for whom we observe child-parent links. This last restriction drops all immigrants and individuals born before 1953 from the sample.

2.2 Definitions

In this section, we discuss our definition of entrepreneurship and details on firm entry, firm ownership, income measurement and the timeline we use in this study. We provide a comprehensive list of key variables and their definitions in Appendix D.

Incorporated and Unincorporated Entrepreneurs. A key question in our study and related earlier work is how to define self-employment and entrepreneurship. The literature has long acknowledged that various types of businesses and their owners can differ significantly from each other in terms of the type of business activity (see e.g. Lazear 2004). For example, Evans and Leighton (1989) and Hurst and Pugsley (2011) illustrate that only a small share of self-employed businesses seek to engage in growth-targeted activities. Levine and Rubinstein (2017) show that incorporated entrepreneurs are more likely to engage in activities that are closely related to tasks that increase productivity and demand stronger cognitive skills, compared to unincorporated business owners, who are less likely to conduct non-routine cognitive tasks. The key legal differences that separate the incorporated from unincorporated firms, namely a separate legal entity and the owners' limited liability, also

make an incorporated business preferable for entrepreneurs who seek to engage in more risky business ventures.

Recent findings by Levine and Rubinstein (2017) and Halvarsson et al. (2018) highlight that the incomes of owners of incorporated and unincorporated businesses differ significantly from each other, such that the owners of incorporated businesses earn much more than other business owners and wage earners in both the US and Sweden. Therefore, pooling all business owners together tends to mask important heterogeneity in income development and income mobility across different types of business owners.

In the Finnish context, there are also several key features that differ between unincorporated firms, consisting of sole proprietors (single owner) and partnership firms, and incorporated firms labeled as privately held corporations. First, the owners of unincorporated firms are personally liable for all the financial commitments such as debts and other obligations of the firm, but the financial liability of incorporated firms is limited to the capital invested in the firm. Second, unincorporated firms are in general subject to more lenient administrative and accounting requirements compared to incorporated firms, including e.g. single-entry bookkeeping, no requirements for capital installments or assigning a board of directors. Third, in the Finnish tax system, unincorporated firms are pass-through entities, meaning that their profits are allocated to their owners as personal taxable income. Incorporated firms constitute separate tax entities, and their profits are taxed at the corporate tax rate. Owners of privately held corporations pay a wage or dividend tax on income withdrawn from the firm.[3] Therefore, as in many other countries and institutional contexts, an unincorporated business tends to be a more suitable choice for those who seek to run a simple business with no immediate intention for growth. In contrast, an incorporated firm is more suitable for entrepreneurs who are oriented towards risk-taking and growth-seeking business activities.

Following these lines of reasoning, we divide business owners into incorporated (IE) and

[3]Appendix C presents a more detailed description of the differences in the tax schedules between the organizational forms.

unincorporated (UIE) entrepreneurs based on the organizational form of the firm in our analysis, allowing us to track down potential differences in income trajectories and mobility outcomes between different types of business owners. We define the organizational form of the business based on its organizational form in the year when it was established, and include the owners in the analysis using this definition even if the firm later changes its organizational form.[4] Our primary focus is on the IE, who we show to be distinctively different from both UIE and wage earners (WE).

Firm Entry. We define the first year of an incorporated business such that it is the first full year after the establishment date when the firm first reports positive sales. We use a similar sales condition for unincorporated businesses but also require the owners to receive at least 50% of their gross market income from the firm in the first year after the establishment date in order to focus on entrepreneurs who are actively engaged in their business operations. This last restriction is important as some wage earners receive additional earnings from an unincorporated business that they run alongside their main occupation, and we do not want to label them as being firm owners in our analysis.

Firm Ownership. Our analysis focuses on the main owners of firms, defined as individuals who own the largest share of the firm when it is established. These individuals typically bear the largest entrepreneurial risk but are also those who are likely to gain most in terms of income if the business ends up being successful. For incorporated business owners, we directly observe the ownership share of the business for each owner in the data, and for owners of unincorporated businesses, we define the main owner status for the owner whose income from the firm is the largest. If two or more individuals own equal shares using the above definitions, they are all considered as main owners. Importantly, we do not condition on the success of the business in our analysis. After the first year, we allow the firm to go

[4]In our data, changing the organizational form is rare, as less than 3% of firms change their organizational form from unincorporated to incorporated within their first 10 years.

bankrupt or the owner to lose the status of main owner, but we still label these individuals as entrepreneurs throughout our sample period.

We narrow our focus to individuals who are *new* owners, defined as individuals who have no current or prior ownership of an incorporated or unincorporated firm in our data. We allow for transitions between the two types of entrepreneurship. This means that individuals can be labeled as new incorporated entrepreneurs if they had no prior ownership of an incorporated business but were sole proprietors or involved in a partnership, and vice versa. We label as wage earners those individuals with positive wage income who have no ownership of *any* firm over the whole period we observe in the data. This allows us to focus on the implications of new business creation on individual income development, excluding those who, for example, join an existing firm as a new partner or inherit an operating business from their parents.[5] Table A1 in Appendix A shows how these definitions and data availability shape the sample size for IE and UIE owners.

We define parental background in entrepreneurship using administrative data starting from 1987. We define the parent(s) of an individual as entrepreneur(s) if he or she is labeled as a business owner for at least 5 years in our data according to the socioeconomic status defined by Statistics Finland. This approach enables us to use a similar definition for parental entrepreneurship using our full data for 1987–2019, as the more detailed ownership data starts only from 1997.

Income Measurement. Our baseline individual income measure is disposable income, which denotes the net-of-tax-and-benefits annual incomes of individuals from all income sources observed in our administrative data. We observe disposable incomes for all wage earners and business owners in each year from 1995 onward, which allows us to follow in-

[5]Previous literature argues that inherited firms constitute only a small share of entrepreneurs. For example, they comprise 5.6% of firms in the US (Parker 2009) and 2.2% in Sweden (Lindquist et al. 2015). Furthermore, our ownership definition excludes those wage earners who, for example, join a firm as a minority partner with a very small ownership share but continue to work in the same firm. These types of ownership arrangements often tend to serve as a commitment tool for high-profile workers and do not represent the type of risky entrepreneurship we aim to focus on in this study.

dividual incomes over a long time period and enables a careful comparison of income development and mobility between the groups. Furthermore, disposable income is a relevant income measure when considering income mobility outcomes in the distribution, as it also captures differences in taxation and benefits both between wage earners and business owners and between the owners of different types of organizational forms. We use market income before taxes and transfers as an alternative income measure, and these results are qualitatively similar to those using disposable income. All income variables are presented in 2005 prices using the Consumer Price Index of Statistics Finland.

In addition to income levels, we analyze the relative income development and income mobility of individuals by using individual and parental income ranks. These denote the individual's relative position in the income distribution by percentile or decile points. We define the rank of each individual in the *full* population consisting of both wage earners and business owners before the firm was established (as explained in more detail below). We define parental income ranks using a distribution of the average annual household market income (labor and capital income) of their parents when they were 45–50 years old. We use market income for the parents to be able to define the ranks of those individuals whose parents were of that age after 1987, as the disposable income measure is not available before 1995. Our results are robust to defining parental ranks using disposable incomes after 1995.

Furthermore, defining an exact boundary between the firm and the owner's incomes can be challenging (see e.g. Kopczuk and Zwick 2020). In addition to income withdrawn from the firm as wages, dividends and capital gains, the owners of privately held corporations can retain earnings in their firm, which do not show up in the administrative data in the year that they are accrued. As an additional analysis to study the relevance of this income source, we follow a similar type of approach as in Alstadsæter et al. (2016) and allocate earnings retained in the firms to each owner of an incorporated firm based on their ownership share of the firm, and include this income on top of the market income observed in the administrative data (see Appendix B for further details).

Timeline. To our knowledge, our study is the first to follow new business owners long before and after they established their first business using population-wide panel data that cover transitions into entrepreneurship over multiple years.[6] This enables us to provide visually clear and tractable evidence on selection into entrepreneurship and the income development and mobility of business owners.

For business owners we define the year $t = 0$ as the year they start their new business, and then follow their income development from before and after they started their business. To obtain a similar timeline for wage earners, we randomly allocate a "pseudo-start year" ($t = 0$) for individuals who over our sample period never engage in any business activities. Therefore, in year $t = 0$ the wage earners do not actually start a business, but this timeline allows us to describe the overall income trajectories of wage earners using a similar timeline as for business owners. This approach provides us with an intuitive benchmark for the general earnings growth in the economy. In our baseline analysis, we follow the income development from 5 years before to 10 years after $t = 0$, i.e., from $t - 5$ to $t + 10$.

In addition to individual income, we use a similar approach to study firm outcomes and how their development differs by income ranks and other characteristics. We focus on six key firm measures, including survival rate, sales, number of employees and labor costs, value added, and the likelihood of R&D investments. As there is no pre-period for new firms, we analyze the development of firm outcomes for their first 10 years, i.e. from $t + 1$ to $t + 10$.

Finally, we acknowledge that starting a business and becoming an entrepreneur are not exogenous events. We further underline this by providing detailed evidence on selection into entrepreneurship below. Therefore, our goal is *not* to provide evidence of the causal effects of starting a business on income development. Instead, we provide visually clear and transparent stylized findings on selection into entrepreneurship, and income development, mobility and firm outcomes for those who self-selected to become entrepreneurs.

[6] Our data include transitions into entrepreneurship that take place between 1998 and 2014. Berglann et al. (2011), the only other study that we know of that follow new entrepreneurs over time in a similar way, focus on transitions that take place between October 2000 and October 2001.

2.3 Descriptive statistics

Panel A of Table 1 presents the cross-sectional descriptive statistics for the full sample in 1997–2019. Similarly to the well-established existing evidence (see e.g. Halvarsson et al. 2018; Levine and Rubinstein 2017), the table illustrates that incorporated entrepreneurs (IE) earn on average much more (mean disposable income 31,097 and market income 40,348 euros) than wage earners (22,353 and 26,995 euros) and unincorporated entrepreneurs (21,565 and 24,679 euros). Also, IEs are more likely to be men. They are only slightly more likely to be highly educated, married, have more children and live in the capital city region compared to wage earners (WE) and unincorporated entrepreneurs (UIE). Also, the groups do not differ by age. Similarly as in many previous studies (Lindquist et al. 2015; Fairlie and Robb 2007; Sorensen 2007; Uusitalo 2001; Dunn and Holtz-Eakin 2000), our data show that both IEs and UIEs are much more likely to have at least one parent who was also an entrepreneur compared to WEs (55%, 61% and 10%, respectively). However, the parental income ranks are rather similar between the groups.

Panel B of Table 1 presents the same statistics for IEs and UIEs one year before they started their business ($t-1$) and for WEs one year before the randomly allocated pseudo-start years. The table shows that those who select to become an IE earn significantly more than WEs and UIEs already *before* they started their business (mean disposable incomes of 28,049, 20,596 and 17,371 euros, respectively). However, the other observed characteristics except sex and parental entrepreneurship are rather similar between the groups, as in the full cross-sectional data.

Furthermore, the majority of the new incorporated business are founded in traditional industries, as shown in Figure A2 in Appendix A. Construction, wholesale, retail trade and restaurants comprise almost 41% of all new firms. Some other notable industries with shares of 4-6% are computer programming, consultancy, and architectural and engineering activities.

3 Selection into Entrepreneurship

In this section, we study selection into new entrepreneurship based on both individual and parental characteristics. We focus on illustrating how the likelihood of starting a new business correlates with individual and parental incomes by showing how the share of new business owners is distributed along individual and parental income ranks. Also, we analyze how individual and parental income ranks are correlated with the initial capital installments of the new business. In addition, we study the role of other observed characteristics in explaining selection into entrepreneurship, such as parental entrepreneurship, age, gender and education.

Selection by individual income. Selection into entrepreneurship by individual income can be conceptualized through three main channels, as formalized by e.g. Levine and Rubinstein (2020): human capital, liquidity constraints and preferences. The human capital or ability channel is often expressed in terms of the jacks-of-all-trades idea popularized by Lazear (2004), which states that entrepreneurs need not be experts in any single skill but rather they need to be multi-skilled generalists. Many of these required skills also have a high value in the labor market, but a decision to start a business requires some premium from entrepreneurship. A large literature has documented that salaried wages and personal wealth have general positive associations with selection into incorporated entrepreneurship (Evans and Jovanovic 1989; Evans and Leighton 1989; Fairlie 1999; Levine and Rubinstein 2017, 2020).

However, correlation between income or wealth and becoming an entrepreneur may also reflect the existence of liquidity constraints. Starting a business typically requires capital, and acquiring these resources externally might be challenging due to, for example, various frictions in the financial market. Therefore, the positive link between income and entrepreneurship can reflect both the human capital channel and the availability of financial resources for starting a business, and it is often difficult to separate between the two in

empirical analysis without an experimental setting.

Our study focuses on selection by income and its broader implications, but preferences for "being your own boss", perceived work flexibility, attitudes toward risk and other similar personality traits are also likely to be correlated with the decision to become an entrepreneur. These types of connections are observed in many studies (see e.g. survey by Kerr et al. 2017). As an example, Hurst and Pugsley (2011) underline that non-pecuniary benefits related to entrepreneurship play a key role in starting a business for many individuals. However, it is not obvious whether or not preferences for entrepreneurship and parental or individual income are linked with each other, making it difficult to disentangle preferences and ability.

We start by zooming into the individual income distribution to analyze self-selection to starting a new business. The upper panel of Figure 1 presents the average likelihood of starting a new incorporated (IE) or unincorporated (UIE) business by individual income rank percentiles, measured one year before individuals started their business ($t-1$). The figure shows the share of new entrepreneurs (vertical axis) coming from each percentile-point in the overall income distribution (horizontal axis), which is measured including both entrepreneurs and wage earners.

From the figure we can observe that there is a clear non-monotonic association between becoming an IE and the individual income rank percentile. Individuals located in the top 20% of the overall income distribution are much more likely to start an incorporated business compared to the rest of the distribution, where the likelihood of starting an incorporated business only weakly increases with income. The share of top income earners is even more striking the closer we zoom into the right tail: more than half of the starters are in the top earnings quartile a year before setting up their first firm, 17% come from the top 5%, and the top 1% spawns more than 5% of new business owners. In contrast, low income individuals are more likely to start an unincorporated business compared to the rest of the distribution, and there is only a weak increase in this likelihood at the very top of the distribution. This means that selection into unincorporated business ownership differs visibly from selection

into incorporated entrepreneurship.

The observed positive connection between income and entrepreneurship and the stark difference in selection between the firm types is well in line with the earlier literature (see e.g. Lazear 2004; Evans and Leighton 1989; Levine and Rubinstein 2017). However, our more granular analysis underlines that the positive selection into new incorporated entrepreneurship appears to hold non-monotonically only at the top of the distribution, instead of continuously increasing by income throughout the distribution. This finding is similar to the non-linear association between individual wealth and starting a business documented previously by Hurst and Lusardi (2004) in the US.

Selection by parental income. The channels regarding selection by parental income are conceptually similar as with individual income. Parental income can be correlated with the human capital and ability of the child, indicating positive selection by parental income into entrepreneurship if these skills can be translated into a positive premium or gain from being a business owner instead of a wage earner. Second, parental income or wealth could be linked to entrepreneurship through the liquidity channel if high-income parents are more able and willing to invest capital in the new business founded by their children. Consistent with these hypotheses, a few papers report a positive relationship between parental income and starting a business (see e.g. Dunn and Holtz-Eakin 2000; Levine and Rubinstein 2017; Lindquist and Vladasel 2022).

The lower panel of Figure 1 illustrates the correlation between parental income rank and the likelihood of becoming a new IE or UIE. In contrast to individual incomes, this likelihood decreases almost linearly with parental income for both types of entrepreneurs, except for the very bottom of the distribution. This means that we do not observe positive selection into any type of new entrepreneurship by parental income, and that the association between individual and parental incomes and selection into entrepreneurship differ significantly from each other, particularly among incorporated entrepreneurs.[7]

[7]Figure A3 in Appendix A shows that this broad observation holds when we use disposable income or

Our observation of the lack of positive selection into entrepreneurship by parental income differs from previous evidence from the US regarding inventors, who are documented to overwhelmingly come from high-income families (Bell et al. 2019). Also, this pattern differs from incorporated business owners in the US and Sweden, where it is observed that family income positively predicts the probability of owning an incorporated business (Dunn and Holtz-Eakin 2000; Levine and Rubinstein 2017; Lindquist and Vladasel 2022).

Our evidence illustrates that these differences between our findings and the previous literature can be explained by our explicit focus on entrepreneurs who start *new* businesses, which departs from previous studies that have focused on entrepreneurship in general – effectively emphasizing entrepreneurs who join an existing firm or whose firms eventually succeeded. First, when using cross-sectional data (years 2014 and 2006) that include all current business owners in Figure 2, we indeed observe that *all* incorporated entrepreneurs more often come from high-income families (upper panel of Figure 2). Second, from the figure we can see a similar pattern when we look at *new* entrepreneurs who join *existing* businesses, as those with higher-income parents are much more likely to join an existing firm as a new owner compared to others. These findings mean that when we focus on mature surviving firms instead of new business creation, we find a positive link between being an incorporated business owner and having high-earning parents. In contrast, using the different samples has practically no effect on the association between parental rank and unincorporated entrepreneurship (lower panel in Figure 2).

One of the factors behind these observations is firm survival. Figure 3 shows that firm survival rates differ notably by entrepreneurs' parental income for IEs. The figure presents the cumulative share of new firms that survived each year over their first 10 years. The upper panel shows the patterns for firm bankruptcies, while the lower panel depicts cases where the individual is no longer the main owner of the business. Bankruptcy is a stringent measure

capital income instead of market income to define parental income ranks. For capital income, roughly a third of new incorporated business owners come from families with zero capital income while the pattern is flat at roughly one percentage per percentile rank from the 40th percentile all the way to the top of the parental capital income distribution.

for a forced firm exit ordered by a court. In comparison, exit from a main ownership status is a more lenient measure, including all channels of exits such as voluntary firm closures, mergers and acquisitions (either partly or fully). The figure shows that firms with owners coming from top income families are much more likely to survive compared to others using both of these measures. This means that due to the higher survival rate, the firms of those with high-income parents tend to be over-represented in samples that include all existing firms, leading to a positive correlation between parental income and being an entrepreneur.

Overall, our observations suggest that parental income is positively linked with individuals joining existing businesses as owners and the likelihood of the survival of the business, but not with the likelihood of starting a completely new firm as a first-time entrepreneur. This evidence brings novel insight to the current understanding of selection into entrepreneurship by highlighting that when looking at new entrepreneurs who establish new businesses, there appears to be no positive gradient between parental income and entrepreneurship. Therefore, the associations between parental income and selection into new entrepreneurship can be clearly different compared to what we can learn from the findings in the previous literature that do not particularly focus on new business creation (Dunn and Holtz-Eakin 2000; Levine and Rubinstein 2017; Lindquist and Vladasel 2022).

Initial equity of the new firm. In order to assess the relevance of the ability or liquidity channels in explaining selection by income, we evaluate the association between the total initial invested equity of new firms and income ranks. We use the equity observed in the firm in its first year to proxy how individual and parental incomes are linked with firms' available capital when they are established.

Figure 4 presents the initial equity of the incorporated firms in logarithmic scale by individual and parental income rank deciles. The figure includes unconditional equity and equity conditional on the industry of the firm, which takes into account the potential differences in the industry composition between the firms in each rank.

The upper graph in Figure 4 shows that there is a clear positive and almost monotonic association between individual incomes at $t-1$ and the equity of the firm in year $t=0$ when the firm is established. This means that individual income is tightly linked with the initial capital assets invested in the firm almost throughout the distribution. In contrast, the lower graph shows that there is no similar connection between parental income and equity, as equity is not positively associated with parental income. This suggests that higher parental income is not reflected in higher capital installments in the firm. One reasonable explanation for this finding is that those who start a new incorporated business are typically older than one might first expect, as the average age of new owners is 35 (see Table 1, Panel B). This suggests that parental income or wealth might not be tightly linked with the financial decisions or actions of their children once they are well into their adulthood. Also, the figure shows that conditioning on industry does not significantly affect any of the equity estimates.

These results imply that even though initial capital assets increase with individual income, they do not explain the observed selection into incorporated entrepreneurship by income as clearly. Above we show that the positive selection by individual income is concentrated only at the very top of the individual distribution (upper panel of Figure 1), whereas initial capital assets increase almost linearly with income. Therefore, this evidence suggests that the large share of top income earners choosing to start a business is at least partly associated with the perceived human capital as an entrepreneur, in contrast to only reflecting liquidity constraints, a finding that echoes the results of Hurst and Lusardi (2004) on individual wealth and starting a business. On the other hand, either ability or liquidity channels do not appear to be directly linked with parental income, as the share of new entrepreneurs (lower panel of Figure 1) and the initial equity of the new firms are not positively correlated with parental income rank.

Selection by other observable factors. Next, we conduct a partial R^2 analysis to simultaneously explore the extent to which various observable characteristics, such as parental entrepreneurship, age and education, explain selection into starting a new incorporated business. These results are reported in Table 2.

First, parental background in entrepreneurship may be associated with both inherited and learned entrepreneurial skills or preferences, or growing into the idea of being entrepreneur (Dunn and Holtz-Eakin 2000; Fairlie and Robb 2007; Lindquist et al. 2015). We find that parental background in entrepreneurship explains up to 75% of the overall observed variation in our model, meaning that this factor very significantly explains selection into entrepreneurship. In our data, more than half of the new IEs and UIEs have at least one entrepreneurial parent (see Table 1). Previous literature has long recognised that parental background in entrepreneurship clearly predicts whether an individual is himself or herself also a business owner (see e.g. Dunn and Holtz-Eakin 2000; Uusitalo 2001; Sorensen 2007; Fairlie and Robb 2007; Lindquist et al. 2015). Our observations clearly demonstrate the importance of intergenerational links in starting a business also in Finland.

However, entrepreneurial background of parents *does not* explain the selection based on individual or parental incomes presented above. Figure 5 presents the shares of new incorporated entrepreneurs by individual (upper graph) and parental (lower graph) income ranks for those with and without entrepreneurial parents. The individual and parental income rank results are strikingly similar for both of these groups, illustrating that this selection pattern does not depend on parental entrepreneurship at all.

Second, Table 2 confirms that the personal income rank at $t-1$ contributes to selection into entrepreneurship by accounting for 7.5% of the total R^2, whereas parental rank does not (0.7% of total R^2). Gender also stands out (5.2%) as most of the new IEs are men (see Table 1). We also find that industry where worked as an employee at $t-1$ plays an important role in explaining selection (9% in total), and in particular individuals working in construction

are more likely to set up an incorporated business (4.3%).[8]

In contrast, age does not significantly explain selection into entrepreneurship. As mentioned above, we find that the owners of new incorporated businesses are seemingly older than one might perhaps first expect. Figure 6 illustrates that new IEs are typically between 30–40 years old (average 35) at the time they start their first business. This point regarding the relatively old age of (successful) business owners has been raised already in the US (Azoulay et al. 2020). In contrast, the largest share of new unincorporated firms are founded by individuals below 30 years of age.[9] Furthermore, being unemployed or studying just before becoming a new entrepreneur appear to matter very little.

The explanatory power of education in Table 2 is very small, meaning that the level of education does not predict starting a new business. However, we find that incorporated entrepreneurs are more likely to have a degree in a technical field such as engineering compared to others, but overall the differences between wage earners and different types of business owners in their level or field of education are small, as illustrated in Figure A4 in Appendix A. Furthermore, Figure A5 in Appendix A shows that the share of highly educated individuals increases linearly with initial income for both types of business owners and wage earners, in contrast to the clear selection into incorporated entrepreneurship by individual initial income taking place only at the top of the distribution.[10]

[8] 95% of new entrepreneurs with their latest employment spell in construction set up their first business in the construction sector. At the other extreme, only 54%, 62% and 66% of entrepreneurs with a background in finance and insurance, manufacturing and wholesale and retail trade, respectively, end up becoming entrepreneurs in the same industry. For other industries, the range is 82–90%.

[9] Note that as we restrict our sample to those individuals for which we observe child-parent links, we drop some of those who start their first business at a relatively old age from our analysis. The average age of starting a new IE would be 39 years in our data without this restriction, in compared to 35 years with this restriction.

[10] Additionally, we illustrate how observable characteristics are correlated with both individual income at $t-1$ and parental income. To do this, we group new entrepreneurs into five individual and five parental income quintiles, and divide the data into 25 bins by these quintiles. Table A2 in Appendix A reveals some notable patterns across the bins. Unsurprisingly, individual income is positively correlated with age, education and gender, while individuals at the bottom of the distribution have more often received unemployment or student benefits in $t-1$. Intriguingly, IEs (UIEs) in the top quintile are three (four) times more likely to have received unemployment benefits than WEs. Note that as we narrow our sample to individuals with positive wage income and/or entrepreneurial income, all of those who received unemployment or student benefits were also active in the labor market in $t-1$. Two stand-out observations emerge when examining individual and parental income distributions jointly. First, IEs in the bottom of both distributions are much less likely

4 Income Mobility

In this section, we first show the average income trajectories of all new business owners around the establishment of their first business. We then focus on the incorporated entrepreneurs in detail to investigate whether entrepreneurship shapes personal incomes differently for individuals coming from different backgrounds.

4.1 Average Income Trajectories

Numerous previous studies suggest that the median business owner does not earn more than their salaried counterpart (see e.g. Evans and Jovanovic 1989, Hamilton 2000, Moskowitz and Vissing-Jørgensen 2002). However, pooling all business owners (self-employed, partnerships, incorporated) together may not be a good proxy for risk-taking entrepreneurship (Glaeser 2009, La Porta and Shleifer 2014). In parallel with this notion, Levine and Rubinstein (2017) disaggregate business owners into entrepreneurs and other business owners based on whether they are incorporated or not. Their results clearly show that the incorporated tend to earn much more than the unincorporated or wage earners. Halvarsson et al. (2018) document a qualitatively similar pattern in Sweden. Based on these studies, however, it is not evident how much incorporated entrepreneurs benefit from entrepreneurship per se and how much of the income differences between the groups can be allotted to self-selection into incorporated entrepreneurship from the top incomes, which we illustrated in the previous section.

To our knowledge, the only previous study that follows new entrepreneurs over time is by Berglann et al. (2011). They focus on transitions into entrepreneurship that take place in Norway during a 12-month spell in 2000–2001. Their findings suggest that incorporated entrepreneurs, in terms of income development, outperform both unincorporated self-employed and the matched control group of wage earners over the first five years of entrepreneurship.

We cover entry into entrepreneurship in 1998–2014, and study income development by

to be unemployed or students compared to UIEs and WEs. Second, IEs in the top of both distributions are slightly less educated than both UIEs and WEs, whereas both IEs and UIEs in this bin are roughly three times as likely to be unemployed as WEs.

analyzing income trajectories five years before and 10 years after the firm was established. To portray the differences in income dynamics between incorporated and unincorporated entrepreneurs and wage earners (pseudo-starters), we estimate the following regression separately for the three groups.

$$y_{ist} = \sum_{j \neq -1} \psi_j \times \mathbb{1}[j = t] + \sum_k \beta_k \times \mathbb{1}[k = s] + X_{is} + \epsilon_{ist}, \tag{1}$$

where y is log income for individual i in year s at event time t, which is the start year of entrepreneurship. We omit the event time dummy at $t-1$ so that the estimated event time coefficients ψ_j represent income development relative to the year just before the start of entrepreneurship and the randomly allocated pseudo-start years for wage earners. β_k represent year dummies flexibly controlling for the overall time trend. In addition to unconditional trajectories, we add various individual-level characteristics (X) to study how they affect the income development patterns of business owners and wage earners. The controls include typical Mincer-type controls (age, gender and level of education, and all their interactions and quadratic and cubic terms for age), and individual and parental percentile income ranks. Individual rank is defined as an average rank before starting a business between the years $t-5$ and $t-1$, and parental income is observed when the parents are 45–50 years old.[11]

The upper panel of Figure 7 shows the unconditional income trajectories and trajectories conditional on Mincer-type controls for both types of business owners (IE and UIE) and

[11]Our empirical approach in studying the income mobility of entrepreneurs has clear similarities with a staggered difference-in-differences setup. As is well known, this approach suffers from challenges related to differences in treatment timing and treatment effect heterogeneity, see Roth et al. (2023) for a recent review. Although our aim is to provide descriptive evidence without causal claims, we want to make sure that the magnitudes of our estimates are not driven by these potential issues. Our setup differs from the examples often discussed in the recent literature, as our benchmark group (wage earners) only consists of "never-treated" individuals. Also, in our setting, the post-entrepreneurship years are always relative to the randomly defined year for the benchmark group, which allows us to estimate the dynamics related to the treatment timing. In Figure 8, we will compare the standard staggered DiD coefficients to the heterogeneity-adjusted coefficients using the doubly-robust DiD estimator from Sant'Anna and Zhao (2020) based on stabilized inverse probability weighting and OLS. Our results using this estimator are similar to our baseline model.

wage earners (WE). The first general finding from the panel is that there is a clear difference between the income development of IEs compared to WEs. The unconditional average income trajectory of IEs is slightly more steeply increasing already before the start of the business, echoing the selection into entrepreneurship based on individual income discussed above. Another interesting finding is the clear dip in the incomes of IE, approximately 8%, in the year the business is established ($t = 0$). This is feasible because it can take some time for a new corporation to be able to operate at full capacity and for the owner to withdraw income from the newly founded firm. The individual incomes of IEs then quickly catch up after this drop and increase more rapidly in the years following the establishment of the business compared to wage earners. After 10 years, entrepreneurs earn on average 21% more than wage earners, relative to the year before the firm was established.

In contrast to IEs, UIEs tend to experience a sharp drop in income one year before they established their business, but then quite rapidly return to a similar stable income trajectory once the business gets operational. The dip at $t - 1$ suggests that entry into UIEs is, on average, driven by different factors than entry into incorporated entrepreneurship. Recall that the average disposable income of UIE is more than a third lower than that of IEs at $t - 1$. Overall, the trajectories illustrate that there are no significant long-run income gains from unincorporated entrepreneurship, and the difference between IEs and UIEs in average income development is very clear.

In addition, the upper panel shows that including Mincer-type controls affects the income trajectories of new business owners only little, but somewhat increases the incomes of WEs. This finding illustrates that controlling for age, gender and education reduces the differences between the income trajectories of IEs and WEs only little (at $t + 10$, from 21% to 17%). The lower panel of Figure 7 shows that when including individual and parental income ranks as additional controls, the trajectories for all three groups flatten substantially. Looking at the difference between IEs and WEs, the trajectories are now even more similar to each other before $t = 0$. The incomes of IEs still increase more rapidly than WEs after the firm

is established, but the difference at $t + 10$ is now further reduced to 16%.[12]

In Figure 8, we narrow our focus to IEs and WEs by plotting the difference in income trajectories between these two groups using various specifications. Although such event study plots are typically used to illustrate dynamic treatment effects, we once again emphasize that our results are descriptive in nature and not the causal effect of entrepreneurship on personal income development. All specifications show the apparent positive trend in income before the start of the business among IEs, the dip in income in the first years of entrepreneurship and the steep increase in income associated with entrepreneurship after the start of the business.

Additionally, in Figure A6 in Appendix A we apply coarsened exact matching (CEM) in an attempt to more carefully balance the observed differences between WE and IE before the firm is established.[13] With CEM we are able to further narrow the differences in the income trends between WEs and IEs before the firm was established. However, the overall implications remain mostly similar as in Figures 7 and 8 above. Significant income gains from incorporated entrepreneurship are evident also when we use matching, as there is an approximately 18% difference in income between IEs and WEs 10 years after the firm was established.

We test the sensitivity of our findings in various dimensions. These results are presented and discussed in more detail in Appendix B. First, focusing on market income instead of disposable income (our baseline measure) does not significantly affect the qualitative aspects

[12]In euros, when looking at averages in the raw data, IEs saw their disposable income increase from 22,000 to 28,000 during the five years before the start of the business, and eventually to 41,000 10 years after becoming entrepreneurs, on average. For WEs, the corresponding numbers are 17,000, 20,000 and 25,000 euros.

[13]In more detail, we use the average income rank deciles between $t-3$ and $t-1$, the growth in rank between $t-3$ and $t-1$, parental income rank, a dummy for a parent being an entrepreneur, average share of labor earnings relative to total income in $t-3$ and $t-1$, and age, gender, place of residence and the number of children at $t-1$ in the CEM weighting procedure to match the groups. For more details on the CEM method, see e.g. Iacus et al. (2012). Applying the CEM method with such a rich set of variables used to balance the groups reduces the sample size considerably. Only 18% of pseudo-starter wage earners and 62% of new incorporated entrepreneurs receive positive CEM weights in the analysis, meaning that we cannot obtain a reasonable comparison using the other observations in the data. This finding further underlines the distinctive self-selection into entrepreneurship based on observed income development before the firm was established when comparing wage earners to incorporated entrepreneurs.

of our analysis. Figure B1 for market income paints a very similar picture as Figure 7 did for disposable income. Second, including earnings retained in the firm and not distributed as wages and dividends for the owners of incorporated firms does not change the overall pattern of income trajectories, but it increases the income levels of IEs by an almost constant annual share of 15% starting right from the first year after the business was established, as illustrated in Figure B1.

Table B1 in Appendix B shows the heterogeneity of the average income trajectories across different subsamples, illustrating that our findings do not crucially change when looking at various subsamples of entrepreneurs. Most importantly, our findings are not sensitive to defining the main owners of incorporated firms using either data from the Finnish Tax Administration or Statistics Finland (discussed in more detail in Section 2). Also, we do not observe significant differences in income gains between entrepreneurs operating in more traditional industries such as manufacturing or transportation, and industries characterized by highly specialized human capital skills, such as consultancy or health services, as illustrated in Figure B3.

4.2 Income Gains by Individual and Parental Income Ranks

Next, we proceed to describe income mobility patterns across the individual and parental income distributions. The average income trajectories presented above leave open the question of whether income developments differ by the backgrounds of individuals. Does entrepreneurship open a way for the individuals at the top of the distribution to become even richer, or does it provide a "Stairway to Heaven" for less well-off individuals to reach the top of the income distribution? Entrepreneurship is often viewed as a key channel for income mobility in society (see e.g. Fairlie 2004), but there is only scarce empirical evidence on this available.

As discussed above in Section 3, personal and parental income may capture differences in innate abilities that can reflect differences in earnings potential and development between entrepreneurs coming from different individual or parental income backgrounds. Also, even

though we argued above that selection into entrepreneurship by income appears not to be fully driven by liquidity, financial constraints may matter for entrepreneurial success as the initial scale of the business can positively impact the success of the business and personal income trajectories (see e.g. Boháček 2006; Quadrini et al. 2009). These channels, therefore, suggest that income development and entrepreneurial success depend positively on personal and parental incomes. However, if the prevalence of these channels is only modest and entrepreneurial success does not crucially hinge on initial monetary resources, the largest gains from entrepreneurship may accrue to those who start with lower income or come from low-income families. This would suggest that entrepreneurship opens up a way to increase individual income particularly for those from less well-off backgrounds.

Furthermore, analyzing the association between individual and parental incomes jointly within the same data and context allows us to evaluate their relative significance for the individual income development of entrepreneurs. Previous literature has studied these associations separately, but it remains an open question whether individual or parental incomes are the main drivers behind individual incomes. To study the distributions jointly, we group IEs and WEs into five individual and parental income quintiles, dividing the data into 25 individual-parental income bins. We then interact these bins with the event timeline in equation (1), and present the differences in income trajectories between IEs and WEs in these groups in Figure 9, in a similar fashion as in Figure 8 above.

Each panel of Figure 9 illustrates the differences in raw income trajectories from $t-5$ to $t+10$ between IEs and WEs by parental income quintiles for a given individual income quintile before $t=0$. The figure suggests that the income gains are rather evenly distributed in all of these groups, except among those who come from high-income families and were already at the top of the individual distribution (the lower-right panel of Figure 9). For these individuals, given their high initial incomes, the gains from entrepreneurship over comparable wage earners translate into over 34,000 euros 10 years after the business was established. In other bins, the monetary increases are 10,000–22,000 euros, with an average of 15,000 eurso

over 10 years.

Figure 9 also reveals that dividing the sample by individual and parent income narrows the differences in income trajectories between entrepreneurs and wage earners prior to the start of the business in most cases. This analysis also depicts that the largest income losses in the year of starting a business are among individuals with the highest original income levels.

Figure A8 in Appendix A summarizes these five-by-five results when including Mincer-type controls, following a similar procedure as in Figure 9. Similarly as above, controlling for age, gender, education and their interactions further narrow the differences in income gains from entrepreneurship by individual and parental incomes, especially in both tails of the income distributions. This figure also shows that the Mincer-controls are important for explaining income trends before entrepreneurship. For the lowest-income quintile, the figure highlights a large drop in income from $t-2$ to $t-1$ and a flat trend before that (see the upper-left panel of Figure A8). Similar pattern is visible for other individual income groups, although to a lesser extent, except that for the richest quintile this approach completely flattens the income trajectory for the five years prior to entrepreneurship (see the lower-right panel of Figure A8).

We also evaluate the importance of various other factors as potential sources of heterogeneity behind the income trajectories within incorporated entrepreneurs and wage earners. We conduct a partial R^2 analysis to study factors explaining income gains between $t-1$ and $t+10$ for IEs and WEs. These results are reported in Table 3.

First, individual income ranks are the most important predictor of income gains for both IEs and WEs – 66% and 76% share of total R^2, respectively. Despite the distinctive selection into incorporated entrepreneurship by those whose parents were also entrepreneurs, parental entrepreneurship seems to matter fairly little (1.6%). This suggests that either inherited ability or entrepreneurial skills learned from parents are not important in explaining the income trajectories of entrepreneurs. Also, age, gender, and industry as an employee at

$t-1$ are not very strongly associated with income gains over time. The same holds for being unemployed or studying at $t-1$. However, highly educated individuals seem to gain somewhat more from entrepreneurship than others. Furthermore, Figure A7 in Appendix A shows that those entrepreneurs who are still running their firms five years after the business was established have larger personal monetary gains, as can be expected. In contrast, those whose firms do not make it past the first five years have similar personal income levels as wage earners 10 years after the business was started.

Overall, our findings demonstrate that entrepreneurship is associated with significant gains in personal income across both personal and parental income distributions. This suggests that entrepreneurship does not only benefit those already at the top, but also provides a "Stairway to Heaven" for those from less well-off backgrounds. However, given the apparent selection into entrepreneurship from the top of the income distribution, it is unclear how entrepreneurship shapes the overall income distribution and top-end inequality. We turn our focus to this issue next.

5 Entrepreneurs at the Top of the Distribution

We tie together our findings on selection and individual income gains to illustrate how entrepreneurship shapes the top of the income distribution. Above we show that top-income individuals are distinctively more likely to establish a new incorporated business than others, but the relative gains from entrepreneurship are rather evenly distributed across both the personal and parental income distributions. This raises the question of what the role of selection is compared to the income gains from entrepreneurship in explaining why business owners tend to govern the top of the income distribution in many developed countries (Smith et al. 2019; Halvarsson et al. 2018; Piketty et al. 2017). Are entrepreneurs over-represented at the top because of the positive associated income gains, or because top income earners are more likely to establish a new business in the first place?

To characterize the role of selection and income gains, we first focus on the top half of the income distribution and describe the share of incorporated entrepreneurs at the top before and after their first firm is established. Figure 10 presents income ranks for IE owners one year before $(t-1)$ and 10 years after $(t+10)$ the firm was established, and income ranks for all incorporated entrepreneurs in the full population of Finnish individuals over 16 years of age in 2006–2016. The latter group thus also includes owners of existing firms that were established earlier.

Figure 10 shows that roughly 5% of incorporated entrepreneurs were among the top 1% of earners already in the year before they established their first business. Ten years later their share increases to almost 11%. Therefore, consistent with our findings on selection and income development, new entrepreneurs appear to occupy the top of the distribution even before starting a business, and their prevalence at the top is clearly increased by the positive income gains associated with incorporated entrepreneurship.

However, it is difficult to pin down whether it is the selection into entrepreneurship or the gains from entrepreneurship that primarily account for the over-representation of entrepreneurs in the top income brackets. Looking at the top 1%, the share of entrepreneurs more than doubles, implying that the two are equal in strength. Looking at the top 5%, the share increases by roughly 50% from 17% to 26%. Such back-of-the-envelope calculations are necessarily incomplete as they do not account for the counterfactual, where the entrepreneurs never established their businesses. Nevertheless, it seems that entrepreneurship amplifies pre-existing differences in income, while the role of selection in explaining the high share of business owners at the top is also large. Furthermore, Figure 10 highlights that the prevalence of *new* incorporated business owners at the top of the distribution 10 years after they established their first firms closely resembles the share of *all* incorporated entrepreneurs at the top. This gives further evidence that the length of our event study window is sufficiently long to capture the over-representation of entrepreneurs at the top of the distribution in the entire population.

To further support the findings portrayed above, Figure 11 presents the likelihood of reaching the top of the income distribution in year $t+10$ for both IEs and WEs, conditional on their income in year $t-1$. The figure shows that IEs are clearly more likely to reach both the top 1% and 10% than WEs throughout the distribution. In particular, the differences are strikingly large among those who come from the bottom half of the distribution, where IEs are 20 times more likely to reach the top 1% compared to WEs. This supports the notion that entrepreneurship can be a "Stairway to Heaven" for those coming from less well-off backgrounds.

In contrast to making it to the top, the bottom graph in Figure 11 illustrates that the likelihood of locating in the bottom 50% of the distribution is rather similar between entrepreneurs and wage earners at all percentile points in the distribution. Overall, Figure 11 thus shows that entrepreneurship is associated with a clear upside in terms of personal income but not with a distinctively increased risk of ending up at the bottom of the distribution. This is consistent with the finding that in addition to significant selection by personal income, incorporated entrepreneurs increase their share at the top of the distribution through notable income gains from entrepreneurship.

We also examine the income composition by income sources to further characterize the role of entrepreneurship at the top of the distribution. In this analysis, we focus on market income instead of disposable income to be able to distinguish between different income sources (wages, business income and other capital income). We evaluate the share of business income of total income across the top 10% of the income distribution, following the work by Smith et al. (2019) for the US.

Figure 12 shows that wage income clearly dominates as an income source from the 90th to the 99th percentile of the overall income distribution of all income earners in Finland. For example, at the 90th percentile approximately 90% of aggregate income consists of wage income. The share of non-wage income (business income + other capital income) gradually increases the closer we approach the very top of the income distribution, and this share

increases very sharply when we move to the richest 0.1%, where the share of wage income falls below 40% of aggregate income, and non-wage income represents more than 60% of total income (solid markers in the figure).

The inclusion of firm-level retained earnings has a notable impact on the income composition at the top. In this analysis we divide earnings retained in the firm and not distributed as wages or dividends to the main owner based on the ownership shares of the firm, as explained in more detail in Section 2.2 and Appendix D. The hollow markers in Figure 12 show that including retained earnings in the market income measure further increases the share of non-wage income at the very top from 60% to almost 80%. This further highlights the role of retained earnings at the very top, similarly as has been recently illustrated using data from the US (Piketty et al. 2017; Kopczuk and Zwick 2020), Norway (Alstadsæter et al. 2016) and the Netherlands (Bruil et al. 2022).

Our results on the income compositions provide very similar evidence as was presented for the US by Smith et al. (2019). Despite the clear differences in the level of overall income inequality, the institutional settings and the tax systems for business owners, business income and other capital income clearly dominate the very top of the distribution in both of these countries. Perhaps surprisingly, the share of non-wage income is even larger for the top 0.1% in Finland compared to the US, especially when retained earnings are accounted for.

6 Firm Performance

Our findings above suggest that the individual-level gains from entrepreneurship are rather evenly distributed across the distributions. However, individual benefits from entrepreneurship do not necessarily go hand in hand with productivity and job creation in the new firms. For example, it could be that high-income individuals are more likely to establish businesses that require few workers but high levels of cognitive skills, such as consultancies, indicating that the firm mainly generates return on the human capital of its main owner. In turn,

middle- and low-income individuals may be more likely to set up businesses that operate in more labor-intensive sectors, implying that the returns are shared more broadly. Therefore, it is crucial to evaluate how firm-level outcomes are associated with owners' background.

Previous literature has argued that young and growing businesses are important for innovation and job creation (see e.g. Decker et al. 2014). Earlier evidence also suggests that the founders of successful businesses are older than commonly thought (Azoulay et al. 2020), innovators often come from the top of the income distribution (Bell et al. 2019) and are responsive to tax incentives (Akcigit et al. 2022), and that education and human capital are important for explaining firm dynamics (Queiro 2022). Our contribution is to analyze how individual-level characteristics – personal and parental income in particular – are connected to key firm outcomes such as employment and productivity.

There is only limited evidence on linking owner characteristics to firm performance measures. However, some notable exceptions exist. Holtz-Eakin et al. (1994) show that receiving an inheritance increases firm survival in the US. Hvide and Møen (2010) find that start-up size increases linearly with an owner's prior wealth from the middle wealth brackets to the top in Norway. Andersen and Nielsen (2012) document that entrepreneurs embarking on a business venture after acquiring windfall wealth have significantly lower survival rates and profits, in contrast to a matched sample of entrepreneurs who did not experience influx of unexpected financial resources in Denmark. Regarding the background characteristics of entrepreneurs more broadly, many studies have focused on the links between entrepreneurs' education and firm success (see e.g. Cooper et al. 1994; Kangasharju and Pekkala 2002; Queiro 2022).

A first important and intuitive measure for the success of a business is its survival rate. Studying firm survival can reveal how successful entrepreneurs coming from different backgrounds are in their business ventures. In Section 3, we briefly discussed the role of survival rates in explaining the differences in the correlation between parental income and entrepreneurship. In Figure 13, we also include the survival rates by individual income, and

illustrate how firm survival rates develop over the first 10 years after the firm was established for owners coming from different individual and parental income deciles.

Panel A of Figure 13 shows that businesses established by top earners face a much smaller likelihood of bankruptcy over their first 10 years. However, this likelihood does not decrease monotonically by individual income, meaning that bottom earners do not establish firms that have the highest chance of being bankrupt early on. Panel B of Figure 13, however, shows that bottom earners have the highest chance of losing majority owner status, meaning that they are either no longer involved in the firm or the firm is no longer active for any other reason. These patterns suggest that firm-level success, measured as survival, is positively linked with entrepreneurs' personal income, although the personal-level income trajectories differ only little across the income distribution, as shown in Section 4.2.

Panel C of Figure 13 shows the bankruptcy rates by parental income deciles. These rates are surprisingly similar between the bottom 90% of the parental income distribution. Instead, firms founded by those who come from the top 10% families face a much smaller likelihood of bankruptcy. This evidence suggests that having high-income parents seems to clearly reduce the risk of a firm going bankrupt. Panel D of Figure 13 shows a similar, although attenuated, pattern regarding majority ownership.

Next, we interact the event time coefficients ψ_j in equation (1) with individual or parental income rank deciles (and their interactions) to study differences in firm success by individual and parental background. We examine six firm-level business activity measures, covering the scale, productivity and innovativeness of the new firms over their first 10 years: annual sales, number of employees and labor costs, value added, value added per employee and the likelihood of positive R&D investments.

Figure 14 presents the results by individual income decile ranks. The main finding from the figure is that key firm-level outcomes appear to be increasing by individual income rank. Firms with owners who had higher incomes before the firm was established grow faster than firms whose main owner had smaller earnings. As an illustration of the magnitudes, the

difference in sales between businesses founded by top and bottom decile earners is approximately 64% (Panel A of Figure 14). For the number of employees (Panel B), staff costs (Panel C), value added (Panel D) and productivity (Panel E), the differences are 43%, 68%, 66% and 33%, respectively. This indicates that the differences are largest in sales, value added and employment. The only exception in Figure 14 is R&D: the pattern in Panel F of the figure suggests a slight U shape on the probability of innovation by income ranks such that the least innovative firms are those with middle-income owners.

Most firm-level outcomes are evenly distributed across the parental income distribution, as shown in Figure 15. This suggests that parental income does not significantly predict the success of the business itself, even though entrepreneurs from high-income families had moderately larger personal income gains than others. The apparent deviation from this pattern is the probability of R&D investments, which is highest for owners from families with the highest income (see Panel F of Figure 15). Therefore, in contrast to individual income, parental income rank does not in itself predict firm success. To further highlight this, in Figure 16 we evaluate how individual and parental income jointly predict the success of new businesses. Similarly as in Figure 14, the figure shows that the scale and productivity of the business increase with individual income even when conditioning on parental income.

Figures 17 and 18 show the development of firm outcomes by individual and parental income separately for entrepreneurs with and without entrepreneurial parents. There are no significant differences by parental entrepreneurship, as the results are very similar between these two groups. If anything, having entrepreneurial parents is weakly positively associated with sales, while the likelihood of R&D investments is higher among business owners with non-entrepreneurial parents. Moreover, firms' success increases with individual income while the patterns are flat over parental income irrespective of whether entrepreneurs have entrepreneurial parents. These findings indicate that despite its significance in selection into entrepreneurship, parental background in entrepreneurship is not tightly associated with either firm-level outcomes or personal income development (see Section 4.2). This is perhaps

surprising in the light of previous literature that underlines the key role of human capital stemming from the entrepreneurial traits of parents (Dunn and Holtz-Eakin 2000; Hurst and Lusardi 2004; Hurst and Pugsley 2011; Lindquist et al. 2015). However, our findings are consistent with Fairlie and Robb (2007), who document that the success of small businesses is only weakly correlated with having a self-employed family member in the US.

We again conduct a partial R^2 analysis to study the relevance of various individual-level characteristics on firm-level success in a unified regression framework. In this analysis we focus on two key firm-level outcomes, the number of employees and value added. Table 4 reveals interesting differences between these two measures. First, entrepreneurs' industry as an employee accounts for a third of the total explanatory power of our model for the number of employees, but only 15% for value added. On average, 82% of new entrepreneurs end up establishing their first business in the same industry as their latest employment spell $(t-1)$ was in. Notably, the restaurant and accommodation sector is positively associated with employment, while the association is negative for manufacturing and architectural and engineering services. For value added, construction stands out with a large negative association.

Also, entrepreneurs' income rank at $t-1$ captures more than a third of the total R^2 for value added, while only 10% for the number of employees. Parental rank (3.3%) and parental entrepreneurship (5.3%) have small but non-negligible shares of R^2 for value added. For employment, their role is very small. With education, we cover 30% of R^2 for employment with a substantial negative association for highly educated individuals, and 19% for value added. Fifth, gender accounts for roughly 10% and 15% for the number of employees and value added, respectively. Finally, age stands out for employment (13%), but not for value added (3.7%).

To summarize, our findings suggest that firms founded by high-income individuals are more successful than those founded by low income earners. This implies that the positive externalities of new entrepreneurship in terms of job creation and productivity are clearly

positively linked with individual income, despite the previous finding that personal income gains from entrepreneurship are not. In addition to entrepreneurial ability, these patterns can partly be explained by liquidity and assets, as the initial equity of the firm also increases with income (see Figure 4).

In contrast to individual incomes, we find no clear association between parental income or parental entrepreneurship and several firm-level outcomes. This suggests that inherited skills, traits or the monetary resources of parents are not very important in determining entrepreneurial success. However, we do observe that the firms founded by those with top income parents are clearly more likely to survive than others. This can indicate, for example, that even though the children of high-income parents do not found more profitable firms, they perhaps engage in less risky business ventures or are more likely to be supported by their parents if their businesses are not successful.

7 Conclusions

In this paper, we provide evidence on selection into entrepreneurship and the income mobility of entrepreneurs. We observe distinctive selection into incorporated entrepreneurship at the very top of the individual income distribution, but no similar selection by parental income. We find that incorporated entrepreneurship is associated with clear positive income gains, compared to both unincorporated business owners and wage earners, and that these gains are quite evenly distributed across both the individual and parental income distributions. Furthermore, we find that new incorporated business owners are much more likely to reach the top income brackets compared to wage earners throughout the pre-entrepreneurship income distribution. Therefore, our findings demonstrate that entrepreneurship does not only benefit those already at the top, but also provides a "Stairway to Heaven" for those from less well-off backgrounds. However, combining our findings on selection and mobility suggests that entrepreneurship amplifies pre-existing top-end income inequality, but the role

of selection into entrepreneurship in explaining the high share of business owners at the top is also considerable.

We show that both the scale (sales, employment) and productivity (value added per worker, R&D investments) of newly established firms increase with individual income. This means that firms founded by top income earners appear to be, on average, more productive and successful than those owned by middle or low-income earners. Thus, they provide the largest positive spillovers in society in terms of job creation and productivity. In contrast, we observe no clear differences in the success of new businesses by parental income. Also, we observe no differences in firm outcomes between those whose parents have an entrepreneurial background and those who do not. These findings reject the popular idea that inherited entrepreneurial traits or human capital are crucial for successful entrepreneurship.

References

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., and Hemous, D. (2018a). Innovation and Top Income Inequality. *The Review of Economic Studies*, 86(1):1–45.

Aghion, P., Akcigit, U., Hyytinen, A., and Toivanen, O. (2018b). On the Returns to Invention within Firms: Evidence from Finland. *AEA Papers and Proceedings*, 108:208–212.

Aghion, P., Akcigit, U., Hyytinen, A., and Toivanen, O. (2022). A Year Older, A Year Wiser (and Farther from Frontier): Invention Rents and Human Capital Depreciation. *The Review of Economics and Statistics*, 104(6):1–31.

Aghion, P., Akcigit, U., Hyytinen, A., and Toivanen, O. (2023). Parental Education and Invention: The Finnish Enigma. *International Economic Review*, 64(2):453–490.

Akcigit, U., Grigsby, J., Nicholas, T., and Stantcheva, S. (2022). Taxation and Innovation in the Twentieth Century. *The Quarterly Journal of Economics*, 137(1):329–385.

Alstadsæter, A., Jacob, M., Kopczuk, W., and Telle, K. (2016). Accounting for Business Income in Measuring Top Income Shares: Integrated Accrual Approach Using Individual and Firm Data from Norway. Working Paper 22888, National Bureau of Economic Research.

Andersen, S. and 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.

Azoulay, P., Jones, B. F., Kim, J. D., and Miranda, J. (2020). Age and High-Growth Entrepreneurship. *American Economic Review: Insights*, 2(1):65–82.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. (2019). Who Becomes an Inventor in America? The Importance of Exposure to Innovation. *The Quarterly Journal of Economics*, 134(2):647–713.

Berglann, H., Moen, E. R., Røed, K., and Skogstrøm, J. F. (2011). Entrepreneurship: Origins and Returns. *Labour economics*, 18(2):180–193.

Boháček, R. (2006). Financial Constraints and Entrepreneurial Investment. *Journal of Monetary Economics*, 53(8):2195–2212.

Bruil, A., van Essen, C., Lejour, A., Möhlmann, J., Rabaté, S., Leenders, W., et al. (2022). Inequality and Redistribution in the Netherlands. Technical report, CPB Netherlands Bureau for Economic Policy Analysis.

Cooper, A. C., Gimeno-Gascon, F. J., and Woo, C. Y. (1994). Initial Human and Financial Capital as Predictors of New Venture Performance. *Journal of Business Venturing*, 9(5):371–395.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The Role of Entrepreneurship in US Job Creation and Economic Dynamism. *Journal of Economic Perspectives*, 28(3):3–24.

Dunn, T. and Holtz-Eakin, D. (2000). Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links. *Journal of Labor Economics*, 18:282–305.

Evans, D. S. and Jovanovic, B. (1989). An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. *Journal of Political Economy*, 97(4):808–827.

Evans, D. S. and Leighton, L. S. (1989). Some Empirical Aspects of Entrepreneurship. *The American Economic Review*, 79(3):519–535.

Fairlie, R. W. (1999). The Absence of African-American Owned Businesses: An Analysis of the Dynamics of Self-Employment. *Journal of Labor Economics*, 17(5):80–108.

Fairlie, R. W. (2004). Does Business Ownership Provide a Source of Upward Mobility for Blacks and Hispanics. In *Public Policy and the Economics of Entrepreneurship*. Cambridge, Mass: MIT Press.

Fairlie, R. W. and Robb, A. (2007). Families, Human Capital, and Small Businesses: Evidence from the Characteristics of Business Owners Survey. *ILR Review*, 60(2):225–245.

Garnadt, N., Funer, L., Stahl, K., and Tåg, J. (2023). Paychecks and Progress: How Labor Market Success Relates to Entrepreneurship Success. Working paper.

Glaeser, E. L. (2009). Entrepreneurship and the City. In *Entrepreneurship and Openness*. Edward Elgar Publishing.

Halvarsson, D., Korpi, M., and Wennberg, K. (2018). Entrepreneurship and Income Inequality. *Journal of Economic Behavior & Organization*, 145:275–293.

Hamilton, B. H. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment. *Journal of Political Economy*, 108(3):604–631.

Harju, J., Koivisto, A., and Matikka, T. (2022). The Effects of Corporate Taxes on Small Firms. *Journal of Public Economics*, 212.

Harju, J. and Matikka, T. (2016). The Elasticity of Taxable Income and Income-Shifting: What is "Real" and What is Not? *International Tax and Public Finance*, 23:640–669.

Holtz-Eakin, D., Joulfaian, D., and Rosen, H. S. (1994). Sticking it Out: Entrepreneurial Survival and Liquidity Constraints. *Journal of Political Economy*, 102(1):53–75.

Hurst, E. and Lusardi, A. (2004). Liquidity Constraints, Household Wealth, and Entrepreneurship. *Journal of Political Economy*, 112(2):319–347.

Hurst, E. and Pugsley, B. W. (2011). What Do Small Businesses Do? *Brookings Papers on Economic Activity*, (43):73–142.

Hvide, H. K. and Møen, J. (2010). Lean and Hungry or Fat and Content? Entrepreneurs' Wealth and Start-Up Performance. *Management Science*, 56(8):1242–1258.

Hyytinen, A., Ilmakunnas, P., and Toivanen, O. (2013). The Return-to-Entrepreneurship Puzzle. *Labour Economics*, 20:57–67.

Hyytinen, A. and Maliranta, M. (2008). When Do Employees Leave Their Job for Entrepreneurship? *Scandinavian Journal of Economics*, 110(1):1–21.

Iacus, S. M., King, G., and Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political analysis*, 20(1):1–24.

Kangasharju, A. and Pekkala, S. (2002). The Role of Education in Self-Employment Success in Finland. *Growth and change*, 33(2):216–237.

Kerr, S. P., Kerr, W. R., and Xu, T. (2017). Personality Traits of Entrepreneurs: A Review of Recent Literature. Working Paper 24097, National Bureau of Economic Research.

Kopczuk, W. and Zwick, E. (2020). Business Incomes at the Top. *Journal of Economic Perspectives*, 34(4):27–51.

La Porta, R. and Shleifer, A. (2014). Informality and Development. *Journal of Economic Perspectives*, 28(3):109–126.

Lazear, E. P. (2004). Balanced Skills and Entrepreneurship. *American Economic Review*, 94(2):208–211.

Levine, R. and Rubinstein, Y. (2017). Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More? *The Quarterly Journal of Economics*, 132(2):963–1018.

Levine, R. and Rubinstein, Y. (2020). Selection Into Entrepreneurship and Self-Employment. CEP discussion paper 1722.

Lindquist, M., Sol, J., and Van Praag, M. (2015). Why Do Entrepreneurial Parents Have Entrepreneurial Children? *Journal of Labor Economics*, 33(2):269–296.

Lindquist, M. and Vladasel, T. (2022). Are Entrepreneurs More Upwardly Mobile? BSE working paper 1351.

Maczulskij, T. and Viinikainen, J. (2023). Self-Confidence Predicts Entrepreneurship and Entrepreneurial Success. *Journal of Business Venturing Insights*, 19.

Maliranta, M. and Nurmi, S. (2019). Business Owners, Employees, and Firm Performance. *Small Business Economics*, 52(1):111–129.

Moskowitz, T. J. and Vissing-Jørgensen, A. (2002). The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle? *American Economic Review*, 92(4):745–778.

Parker, S. (2009). *The Economics of Entrepreneurship*. Cambridge University Press, Cambridge.

Paukkeri, T., Ravaska, T., and Riihelä, M. (2023). The role of privately held firms in income inequality. IFS working paper 23/36.

Piketty, T., Saez, E., and Zucman, G. (2017). Distributional National Accounts: Methods and Estimates for the United States. *The Quarterly Journal of Economics*, 133(2):553–609.

Quadrini, V. et al. (2009). Entrepreneurship in Macroeconomics. *Annals of Finance*, 5(3):295.

Queiro, F. (2022). Entrepreneurial Human Capital and Firm Dynamics. *The Review of Economic Studies*, 89(4):2061–2100.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. *Journal of Econometrics*, 235(2):2218–2244.

Sant'Anna, P. H. and Zhao, J. (2020). Doubly Robust Difference-in-Differences Estimators. *Journal of Econometrics*, 219(1):101–122.

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

Smith, M., Zidar, O., and Zwick, E. (2022). Top Wealth in America: New Estimates Under Heterogeneous Returs. *The Quarterly Journal of Economics*, Forthcoming.

Sorensen, J. (2007). Closure and Exposure: Mechanisms in the Intergenerational Transmission of Self-Employment. *Research in the Sociology of Organizations*, 25:83–124.

Uusitalo, R. (2001). Homo Entreprenaurus? *Applied economics*, 33(13):1631–1638.

Tables and Figures

Table 1: Descriptive statistics

			Entrepreneurs		
	All	Wage earners	All	Unincorporated	Incorporated
Panel A: Full sample, 1997-2019					
Observations	9,214,232	7,994,448	1,219,784	866,905	403,565
Share of observations	100%	86.5%	13.5%	9.6%	4.5%
Disposable income	22,637	22,353	24,455	21,565	31,097
Disposable income rank	51	52	50	45	61
Market income	27,313	26,995	29,393	24,679	40,348
Market income rank	51	51	49	44	61
Parental income rank	35	34	37	36	40
Parent entrepreneur	0.16	0.10	0.58	0.61	0.55
Age	37	37	36	36	37
Female	0.48	0.51	0.33	0.37	0.24
Secondary education or higher	0.42	0.44	0.34	0.30	0.42
Tertiary education or higher	0.28	0.28	0.23	0.20	0.30
Capital city region	0.22	0.22	0.23	0.22	0.25
Married	0.47	0.47	0.48	0.46	0.54
Number of children	1.13	1.11	1.21	1.18	1.29
Panel B: Sample at *t-1*, 1997-2013					
Observations	659,446	569,044	90,402	61,876	28,707
Share of observations	100%	86.0%	14.0%	9.5%	4.5%
Disposable income	20,619	20,596	20,765	17,371	28,049
Disposable income rank	49	50	43	36	59
Market income	24,424	24,499	23,954	18,081	36,623
Market income rank	49	50	43	35	59
Parental income rank	35	34	37	36	40
Parent entrepreneur	0.16	0.10	0.58	0.61	0.55
Age	34	34	34	33	35
Female	0.49	0.51	0.33	0.38	0.25
Secondary education or higher	0.41	0.43	0.34	0.30	0.43
Tertiary education or higher	0.27	0.27	0.23	0.20	0.30
Capital city region	0.23	0.23	0.23	0.23	0.25
Married	0.43	0.42	0.45	0.42	0.51
Number of children	1.06	1.04	1.16	1.12	1.25

Notes: Table presents the mean value of each variable for all individuals, wage earners and owners of unincorporated and incorporated businesses. Panel A presents the statistics for the full sample in 1997–2018, including all wage earners and business owners. Panel B describes the sample in year $t-1$, one year before the business was established and one year before the randomly allocated pseudo-starts for wage earners for the years 1998–2013. This sample includes the main owners of businesses, and is limited to the year 2013 as we follow each individual for a minimum of five years. Secondary education includes high school and vocational education. Tertiary education includes a bachelor's or a higher degree. Capital city region includes Helsinki, Espoo, Vantaa and Kauniainen. The detailed definitions of the variables are presented in Appendix D.

Table 2: Probability of starting an incorporated business by observable characteristics: Partial R^2 analysis

	Partial coeff.	Partial R^2	Share of total R^2
Individual rank at $t-1$	-0.0558	0.0031	2.62%
(Individual income rank)^2	0.0762	0.0058	4.88%
Parental income rank	0.0288	0.0008	0.70%
(Parental income rank)^2	-0.0005	0.0000	0.00%
Parent entrepreneur	0.2980	0.0888	74.69%
Industry as an employer			
Manufacturing	-0.0400	0.0016	1.34%
Construction	0.0714	0.0051	4.29%
Wholesale and retail trade	0.0373	0.0014	1.17%
Restaurants and accommodation	0.0247	0.0006	0.51%
Finance and insurance	-0.0106	0.0001	0.09%
Legal and accounting services	0.0199	0.0004	0.33%
Management consulting	0.0326	0.0011	0.89%
Architectural and engineering services	0.0199	0.0004	0.33%
Health services	0.0068	0.0000	0.04%
Total		0.0107	9.01%
Age at $t-1$	0.0208	0.0004	0.36%
Age^2	-0.0190	0.0004	0.30%
Education			
Secondary or higher	-0.0071	0.0001	0.04%
Tertiary or higher	-0.0049	0.0000	0.02%
Unknown	0.0130	0.0002	0.14%
Female	-0.0785	0.0062	5.19%
Unemployed at $t-1$	0.0471	0.0022	1.86%
Student at $t-1$	-0.0149	0.0002	0.19%
Total R^2		0.1189	100%

Notes: Table presents the estimates of the proportion of the variance of the probability of starting an incorporated business that is explained by each variable. The estimates are based on a sample of those individuals who establish an incorporated business and those who never engage in any business activities (wage earners). Partial coeff. refers to the partial correlation coefficient of a specified variable after removing the effects of all other variables in the variable list, and Partial R^2 denotes the square of this coefficient. Share of total R^2 denotes how much the specified variable explains the total R^2 of the model. Secondary education includes high school and vocational education. Tertiary education includes a bachelor's or a higher degree. The detailed definitions of the variables are presented in Appendix D.

Table 3: Gains from entrepreneurship from $t-1$ to $t+10$ by observable characteristics: Partial R^2 analysis

	Incorporated entrepreneurs			Wage earners		
	Partial coeff.	Partial R^2	Share of total R^2	Partial coeff.	Partial R^2	Share of total R^2
Individual income rank at $t-1$	-0.1931	0.0373	48.47%	-0.2730	0.0745	48.52%
(Individual income rank)^2	0.1168	0.0136	17.74%	0.2052	0.0421	27.42%
Parental income rank	-0.0083	0.0001	0.09%	-0.0221	0.0005	0.32%
(Parental income rank)^2	0.0170	0.0003	0.38%	0.0208	0.0004	0.28%
Parent entrepreneur	0.0346	0.0012	1.56%	-0.0133	0.0002	0.12%
Industry as an employee						
Manufacturing	-0.0171	0.0003	0.38%	-0.0375	0.0014	0.91%
Construction	-0.0304	0.0009	1.20%	-0.0209	0.0004	0.28%
Wholesale and retail trade	-0.0120	0.0001	0.19%	-0.0287	0.0008	0.54%
Restaurants and accommodation	-0.0238	0.0006	0.73%	-0.0275	0.0008	0.49%
Finance and insurance	-0.0033	0.0000	0.01%	0.0067	0.0000	0.03%
Legal and accounting services	0.0020	0.0000	0.01%	-0.0043	0.0000	0.01%
Management consulting	-0.0013	0.0000	0.00%	-0.0025	0.0000	0.00%
Architectural and engineering services	0.0076	0.0001	0.08%	-0.0095	0.0001	0.06%
Health services	0.0039	0.0000	0.02%	-0.0268	0.0007	0.47%
Total		0.0020	2.62%		0.0043	2.80%
Age at $t-1$	0.0025	0.0000	0.01%	0.0157	0.0002	0.16%
Age^2	-0.0119	0.0001	0.18%	-0.0319	0.0010	0.66%
Education						
Secondary or higher	0.0449	0.0020	2.62%	0.0748	0.0056	3.64%
Tertiary or higher	0.1090	0.0119	15.44%	0.0956	0.0091	5.95%
Unknown	-0.0269	0.0007	0.94%	-0.0337	0.0011	0.74%
Female	-0.0665	0.0044	5.75%	-0.0683	0.0047	3.03%
Unemployed at $t-1$	-0.0477	0.0023	2.96%	-0.0505	0.0026	1.66%
Student at $t-1$	0.0308	0.0009	1.23%	0.0849	0.0072	4.69%
Total R^2		0.0769	100%		0.1536	100%

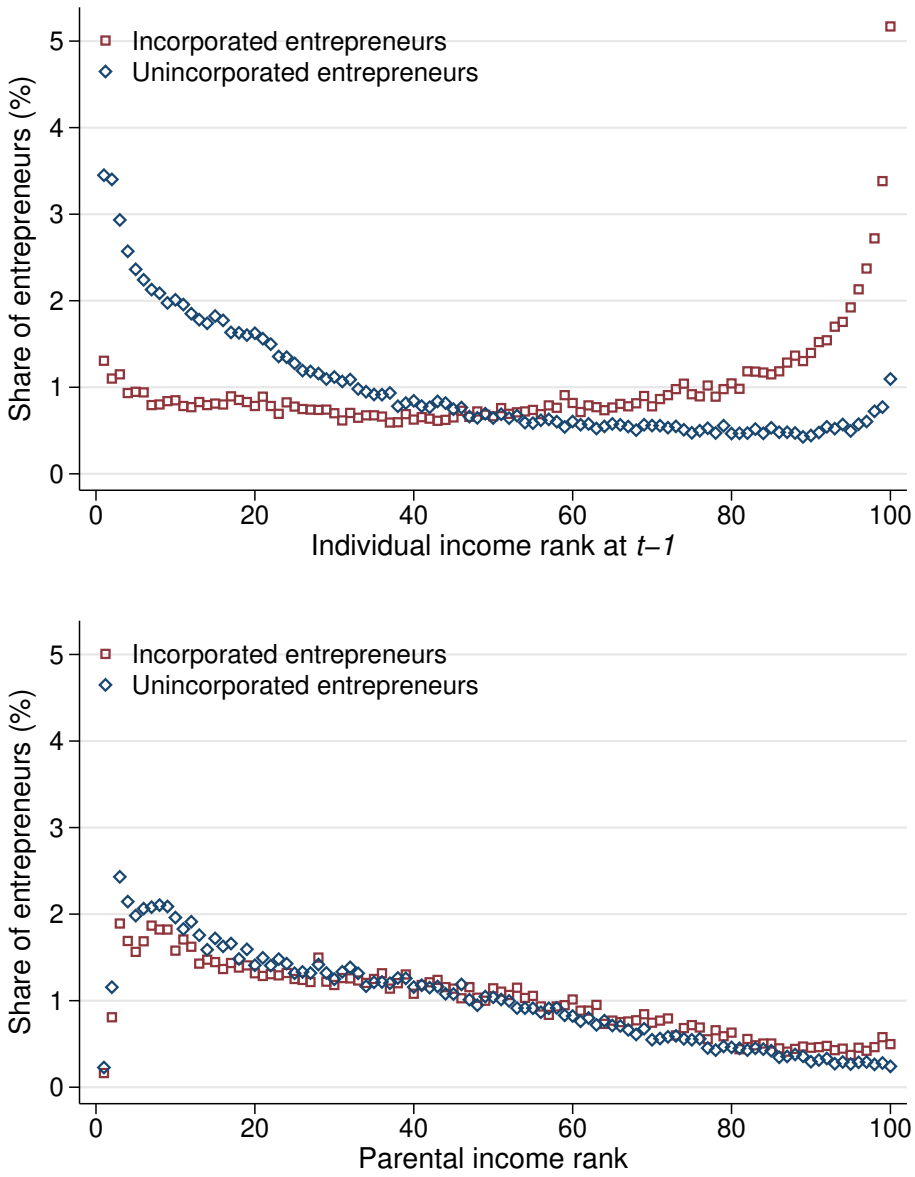
Notes: Table presents the estimates of the proportion of the variance of the increase in disposable income from $t-1$ to $t+10$ that is explained by each predictor. For incorporated entrepreneurs, $t=0$ indicates the year when their first business was established. For wage earners, $t=0$ represents the randomly allocated pseudo-start. The estimates are based on separate regressions for entrepreneurs and wage earners. Partial coeff. refers to the partial correlation coefficient of a specified variable after removing the effects of all other variables in the variable list, and Partial R^2 denotes the square of this coefficient. Share of total R^2 denotes how much the specified variable explains the total R^2 of the model. Secondary education includes high school and vocational education. Tertiary education includes a bachelor's or a higher degree. The detailed definitions of the variables are presented in Appendix D.

Table 4: Firm-level measures by observable individual-level characteristics: Partial R^2 analysis

	Number of employees			Value added		
	Partial coeff.	Partial R^2	Share of total R^2	Partial coeff.	Partial R^2	Share of total R^2
Individual rank at $t-1$	0.0526	0.0028	10.05%	0.0921	0.0085	25.79%
(Individual rank)^2	-0.0034	0.0000	0.04%	-0.0449	0.0020	6.14%
Parental rank	0.0122	0.0002	0.55%	0.0253	0.0006	1.94%
(Parental rank)^2	-0.0072	0.0001	0.19%	-0.0207	0.0004	1.30%
Parent entrepreneur	0.0093	0.0001	0.32%	0.0418	0.0017	5.32%
Industry as an employee						
Manufacturing	-0.0382	0.0015	5.30%	0.0082	0.0001	0.21%
Construction	-0.0266	0.0007	2.58%	-0.0516	0.0027	8.09%
Wholesale and retail trade	0.0190	0.0004	1.32%	0.0285	0.0008	2.48%
Restaurants and accommodation	0.0479	0.0023	8.35%	0.0127	0.0002	0.49%
Finance and insurance	-0.0449	0.0020	7.33%	-0.0188	0.0004	1.08%
Legal and accounting services	0.0019	0.0000	0.01%	0.0096	0.0001	0.28%
Management consulting	0.0224	0.0005	1.82%	0.0159	0.0003	0.77%
Architectural and engineering services	-0.0400	0.0016	5.81%	-0.0191	0.0004	1.11%
Health services	-0.0123	0.0002	0.55%	0.0044	0.0000	0.06%
Total		0.0091	33.07%		0.0048	14.56%
Age at $t-1$	0.0352	0.0012	4.51%	0.0190	0.0004	1.10%
Age^2	-0.0496	0.0025	8.94%	-0.0292	0.0009	2.60%
Education						
Secondary or higher	0.0480	0.0023	8.39%	0.0482	0.0023	7.07%
Tertiary or higher	-0.0627	0.0039	14.30%	-0.0284	0.0008	2.45%
Unknown	0.0459	0.0021	7.65%	0.0565	0.0032	9.71%
Female	-0.0534	0.0028	10.34%	-0.0712	0.0051	15.44%
Unemployed at $t-1$	-0.0005	0.0000	0.00%	0.0271	0.0007	2.23%
Student at $t-1$	0.0213	0.0005	1.65%	0.0378	0.0014	4.35%
Total R^2		0.0275	100%		0.0329	100%

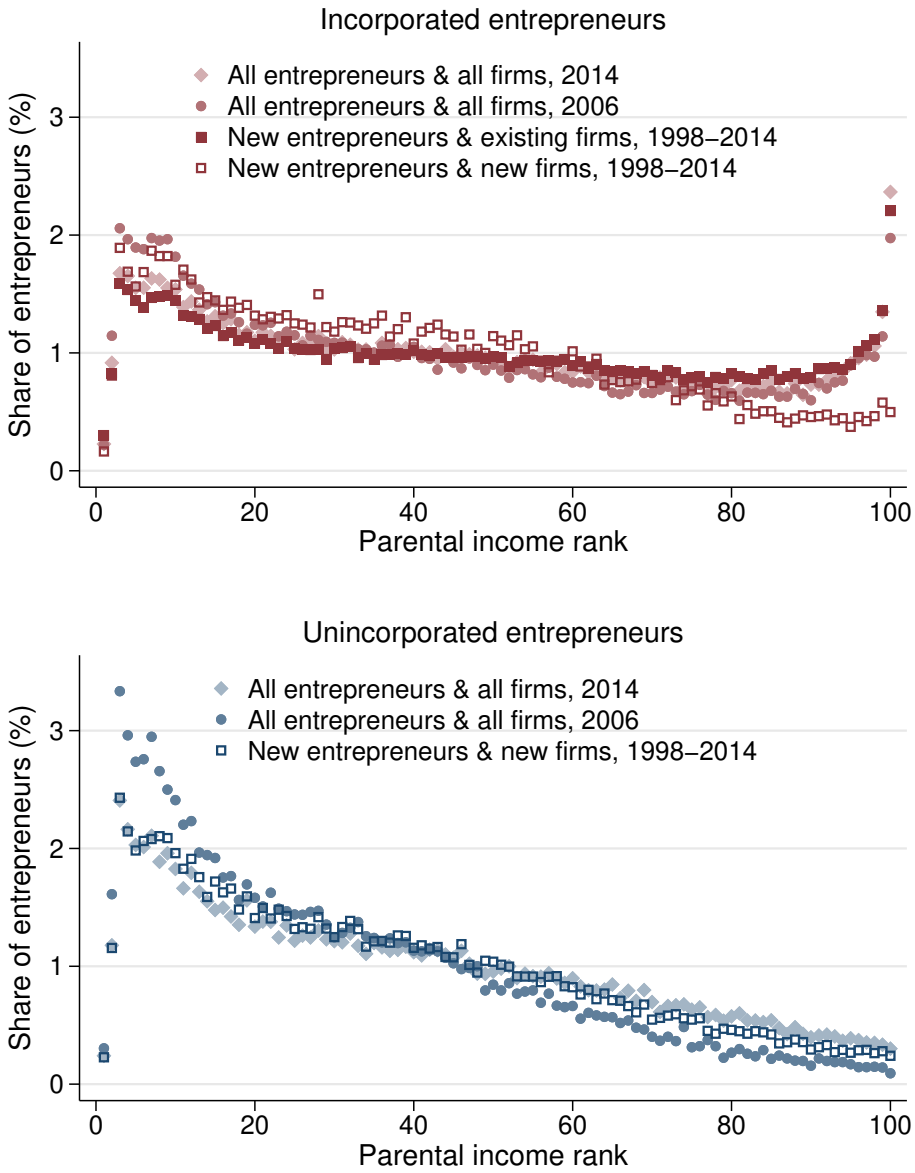
Notes: Table presents the estimates of the proportion of the variance of firm-level measures that is explained by each individual-level predictor. The two measures, number of employees and value added, are observed 10 years after the firm was established including those incorporated businesses that are still active 10 years after $(t+10)$. Partial coeff. refers to the partial correlation coefficient of a specified variable after removing the effects of all other variables in the variable list, and Partial R^2 denotes the square of this coefficient. Share of total R^2 denotes how much the specified variable explains the total R^2 of the model. Secondary education includes high school and vocational education. Tertiary education includes a bachelor's or a higher degree. The detailed definitions of the variables are presented in Appendix D.

Figure 1: Share of new entrepreneurs by individual and parental income ranks



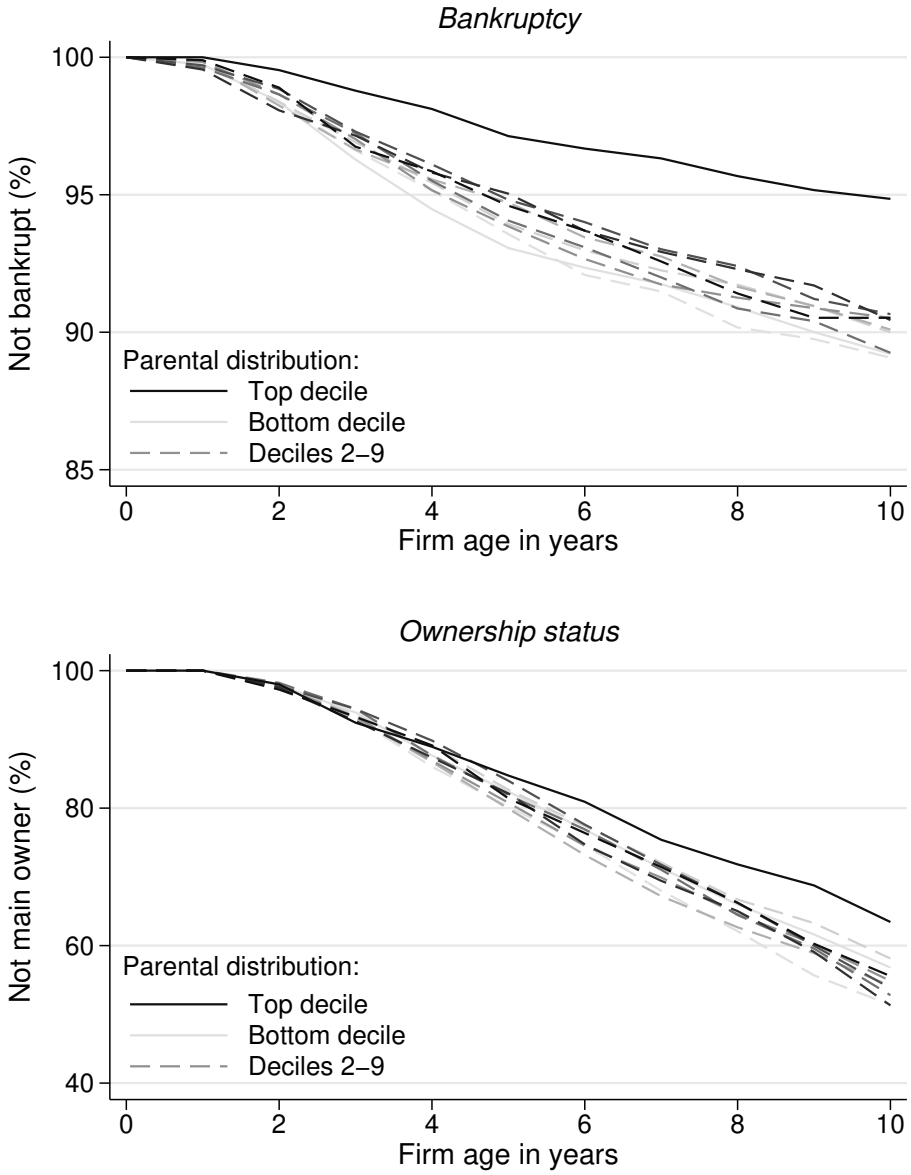
Notes: Figure presents the percentage shares of new incorporated (red squares) and unincorporated (blue diamonds) business owners in each individual income rank percentile (upper graph) measured one year before establishing the business ($t-1$), and by their parental income rank percentile (bottom graph). Individual income ranks are calculated from the distribution of disposable income in $t-1$ including all individuals (entrepreneurs and wage earners) in our baseline sample. Parental income is calculated as average annual household market income when the parents were 45–50 years old. The figure shows that a large share of incorporated entrepreneurs come from the top of the income distribution, but there is no similar selection for unincorporated entrepreneurs. For example, the figure shows that slightly more than 5% of new incorporated entrepreneurs were among the top 1% of all income earners one year before setting up their first incorporated business. In contrast, the likelihood of starting a new business decreases with parental income rank.

Figure 2: Share of entrepreneurs by parental income: New firms vs. all existing firms



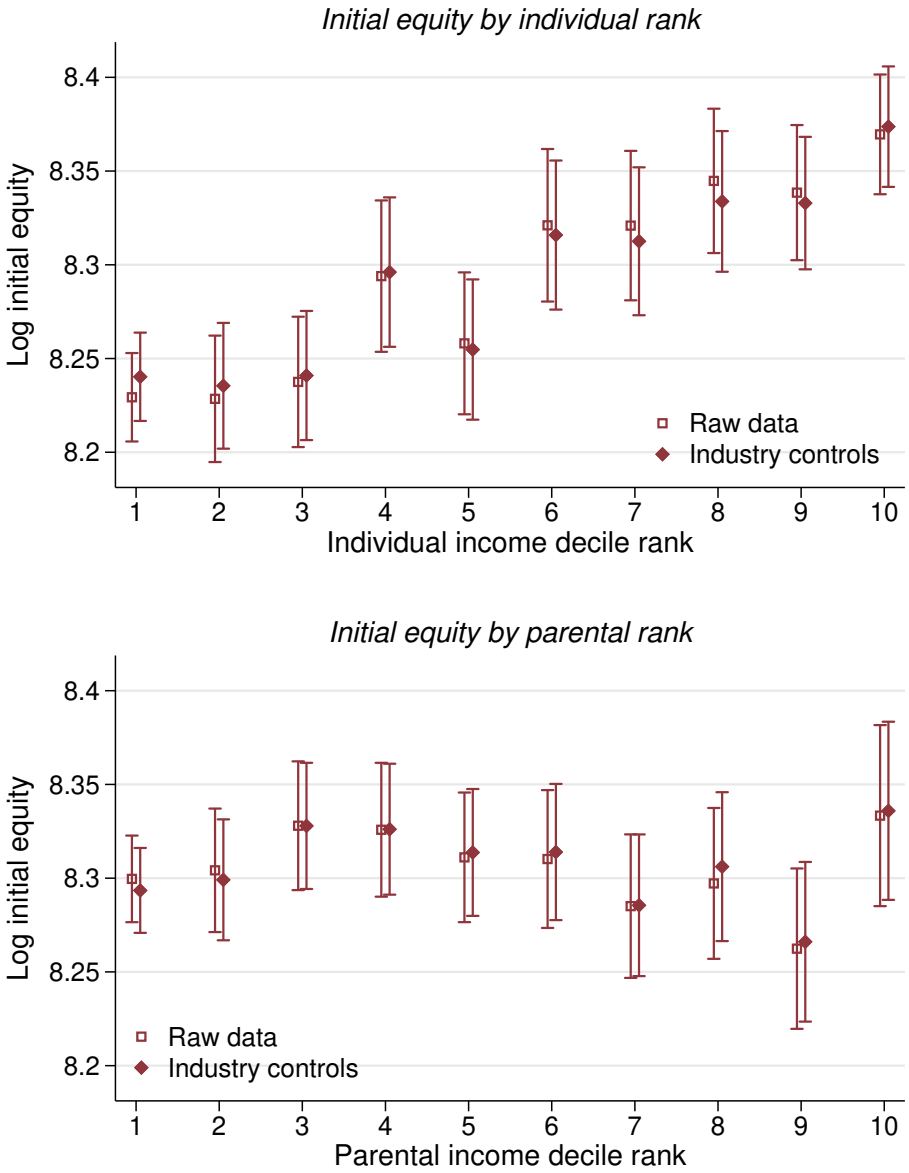
Notes: Figure presents the percentage shares of incorporated (upper graph) and unincorporated (bottom graph) business owners by their parental income ranks using four different definitions of entrepreneurship. The hollow square markers correspond to the lower panel of Figure 1 and show the shares of new business owners using data for 1998–2014. The circle and diamond markers show the shares for all existing entrepreneurs using two cross-sections of data (2006 and 2014). The red square markers show the figures for new entrepreneurs who join existing firms using data for 1998–2014. Parental income is calculated as average annual household market income when the parents were 45–50 years old. The figure shows that for existing firms there is a positive link for parental income and entrepreneurship (as in previous literature), but this link does not exist for new incorporated firms and new business creation.

Figure 3: Survival rates of new incorporated businesses by parental income ranks



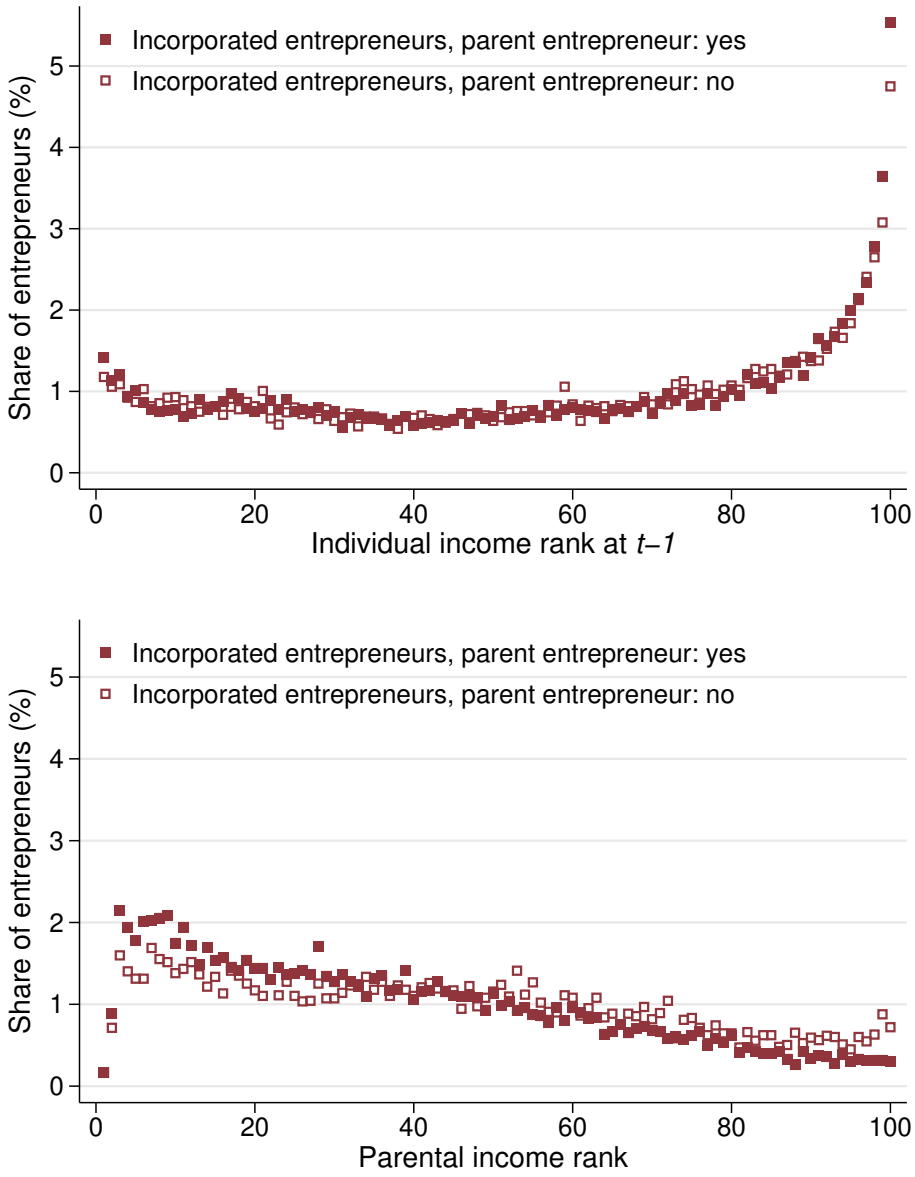
Notes: The upper panel of the figure shows the share of incorporated businesses that are not bankrupt in each firm age year (1–10) by the parental income ranks of the main owner. The top parental income rank decile is marked with a solid black line, and the bottom decile with a solid grey line. Deciles 2–9 are marked with dashed lines. The lower panel shows a similar survival rate graph by the main owner status of the business, indicating the share of new firms with the same initial main owner still having main ownership in the firm. Parental income is calculated as average annual household market income when the parents were 45–50 years old. The figure shows that the survival rates are higher using both measures among those with highest parental incomes.

Figure 4: Initial equity of new incorporated businesses by individual and parental income ranks



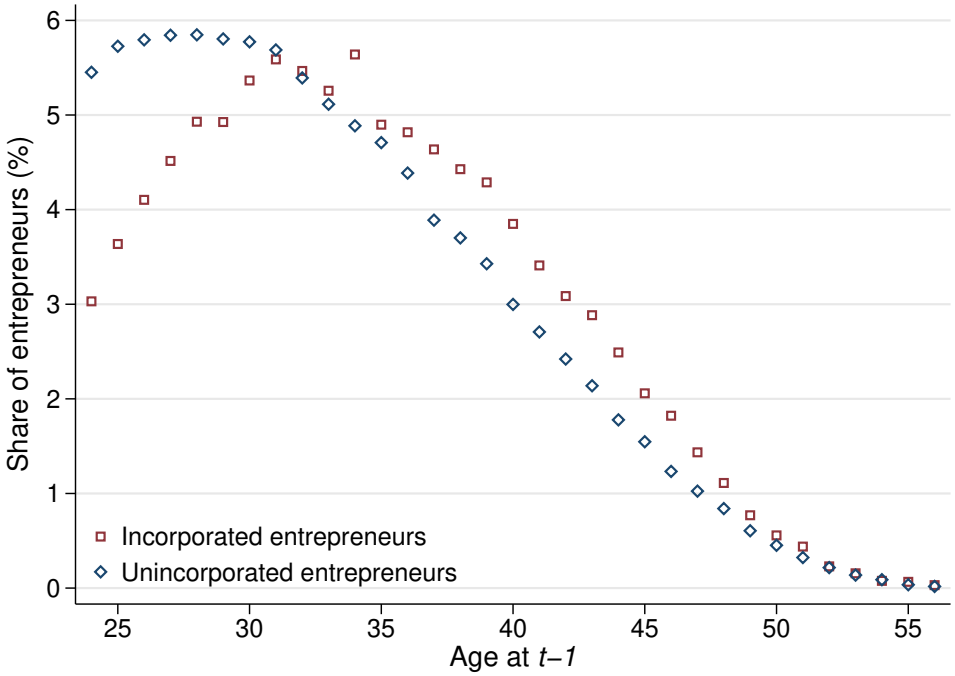
Notes: The figure shows the average total initial invested equity of incorporated businesses in logs in the year of establishment by the owners' individual income ranks and by their parental income ranks with 95% confidence intervals. Individual income ranks are calculated as an average over the years $t-5$ and $t-1$ including all individuals in our baseline sample. Parental income is calculated as average annual household market income when the parents were 45–50 years old. The figure includes unconditional equity marked with hollow squares and equity conditional on the industry of the firm, marked with red diamonds. The figure shows that initial equity increases linearly with individual income, but there is no association between new firms' initial equity and parental income. Controlling for the industry of the new firm does not affect the estimates.

Figure 5: Share of new incorporated entrepreneurs by parental entrepreneurship



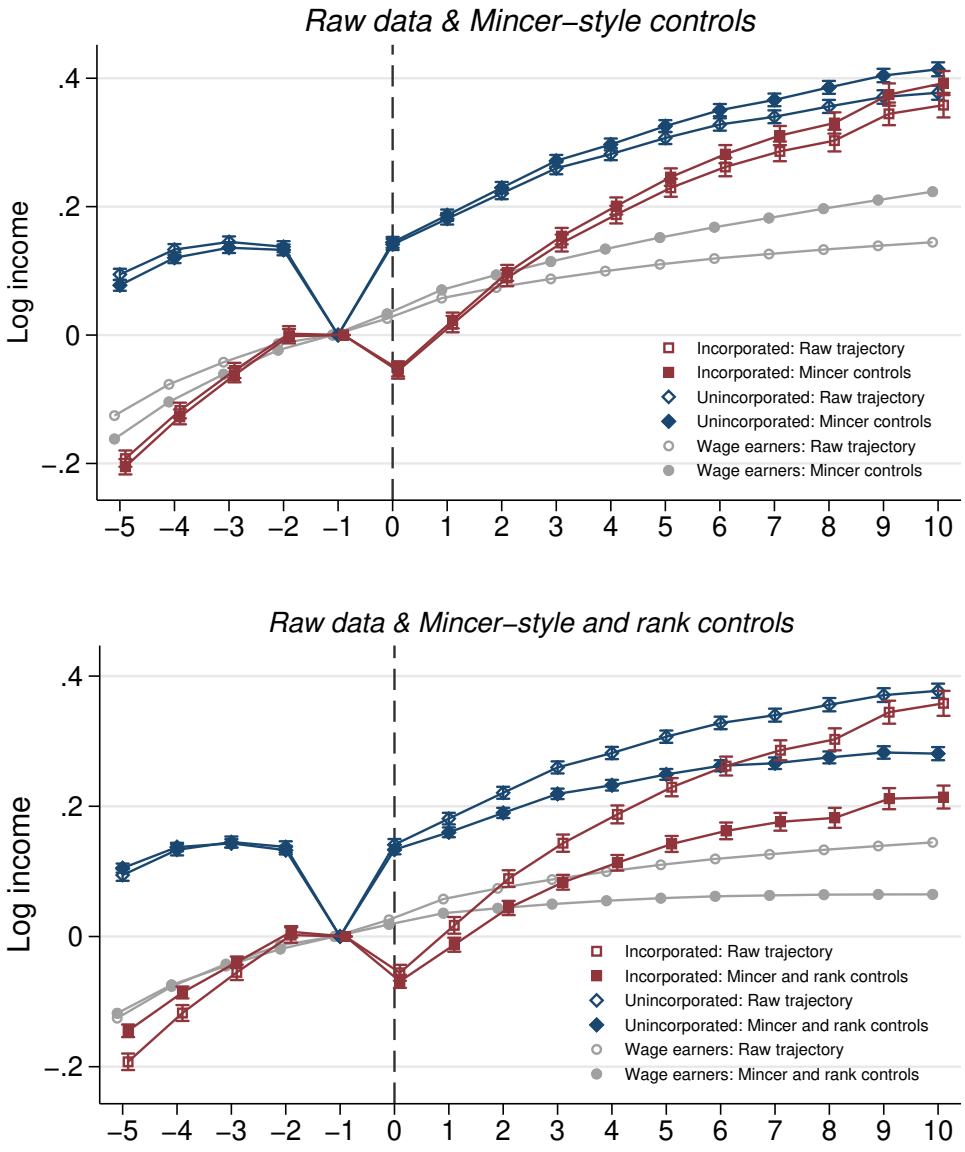
Notes: Figure presents the percentage shares of new incorporated business owners in each individual income rank percentile (upper graph) measured one year before establishing the new business ($t-1$) and by their parental income rank percentiles (bottom graph). The sample of entrepreneurs is divided in two depending whether at least one of their parents was an entrepreneur. Individual income ranks are calculated from the distribution of disposable income in $t-1$ including all individuals in our baseline sample. Parental income is calculated as average annual household market income when the parents were 45–50 years old. Parents are classified as entrepreneurs if they are labeled as a business owner for at least 5 years according to the socio-economic status defined by Statistics Finland. The figure shows that the shares of entrepreneurs in each income rank do not significantly differ between those with and without entrepreneurial parents.

Figure 6: Share of new entrepreneurs by age



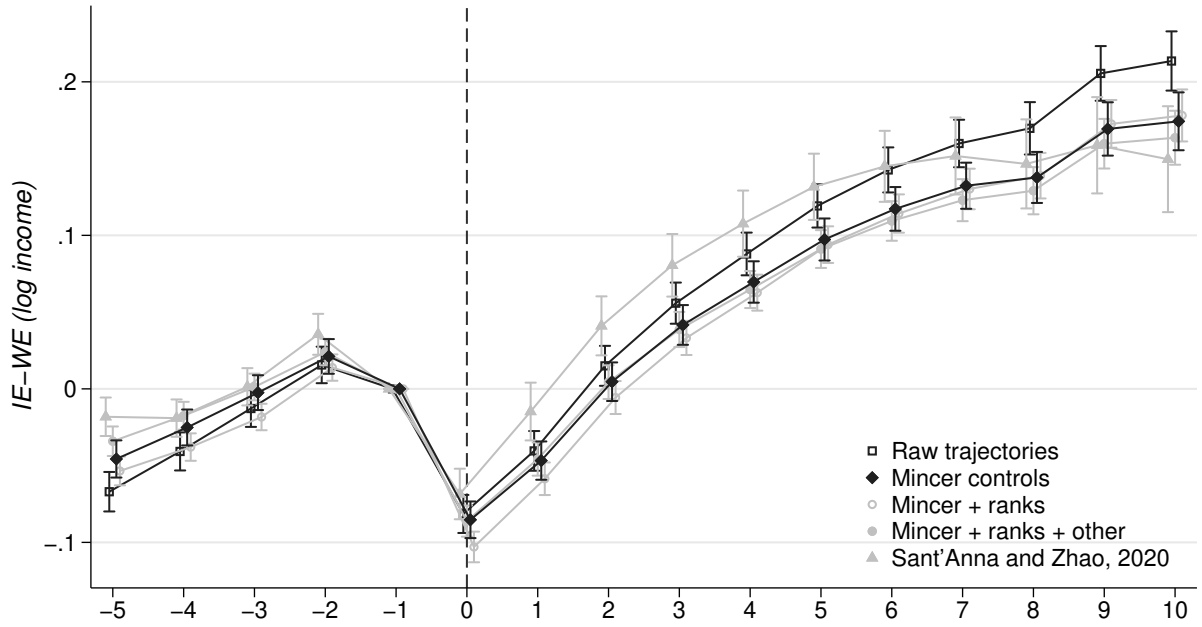
Notes: Figure presents the percentage shares of new incorporated (red squares) and unincorporated (blue diamonds) business owners by their age one year before establishing the business ($t-1$). The average age of new incorporated and unincorporated business owners in our baseline sample is 35 and 33 years, respectively.

Figure 7: Average income trajectories



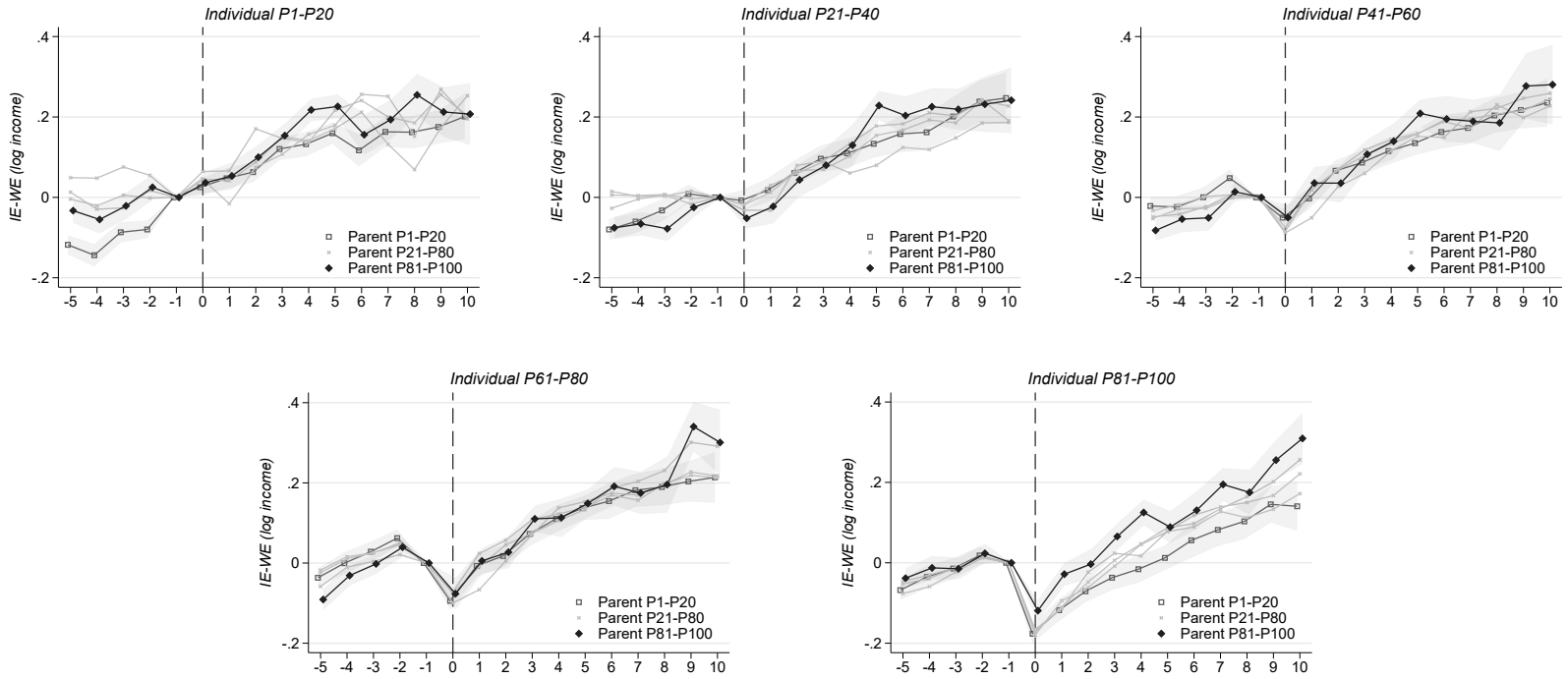
Notes: The figure presents the log disposable income trajectories of new business owners (incorporated and unincorporated) and wage earners before and after the business was established ($t = 0$), relative to one year before the firm was started ($t - 1$), denoted by zero in the figure with 95% confidence intervals. For wage earners, pseudo-starts at ($t = 0$) are randomly drawn from a uniform distribution. Raw trajectories correspond to equation (1) without any additional control variables. The Mincer-style controls include all interactions of age, sex and education (primary education, secondary degree or tertiary degree) and the interactions using age squared and age cubed. In the lower panel, raw trajectories are compared with trajectories when including both Mincer-style controls and individual and parental income ranks as controls. Individual income ranks are measured as an average over the years $t - 5$ and $t - 1$ from the full distribution including both wage earners and business owners in our baseline sample. Parental income ranks are calculated using the average of annual household income when the parents were 45–50 years old.

Figure 8: Difference in income trajectories between incorporated entrepreneurs and wage earners



Notes: The figure presents the difference in log disposable income trajectories between incorporated entrepreneurs (IE) and wage earners (WE) before and after the business was established ($t = 0$), relative to the year $t - 1$ (denoted by zero in the figure) with 95% confidence intervals using different estimation specifications. For wage earners, pseudo-starts are drawn from a uniform distribution. The Mincer-style controls include all interactions of age, sex and education (primary education, secondary degree or tertiary degree) and the interactions using age squared and age cubed. Individual income ranks are measured as an average over the years $t - 5$ and $t - 1$ from the full distribution including both wage earners and business owners in our baseline sample. Parental income ranks are calculated using the average of annual household income when the parents were 45–50 years old. The additional controls (other) include industry as an employer at $t - 1$, marital status, number of children, place of residence, and information on unemployment and student benefits, social assistance and housing allowances and parental allowances in $t - 1$. Sant'Anna and Zhao, 2020 refer to the doubly-robust difference-in-differences estimator from Sant'Anna and Zhao (2020) based on stabilized inverse probability weighting and OLS with no control variables. The figure shows that the trajectories lead to very similar outcomes after 10 years ($t + 10$) regardless of controls or estimators.

Figure 9: Differences in income trajectories between incorporated entrepreneurs and wage earners by individual and parental income

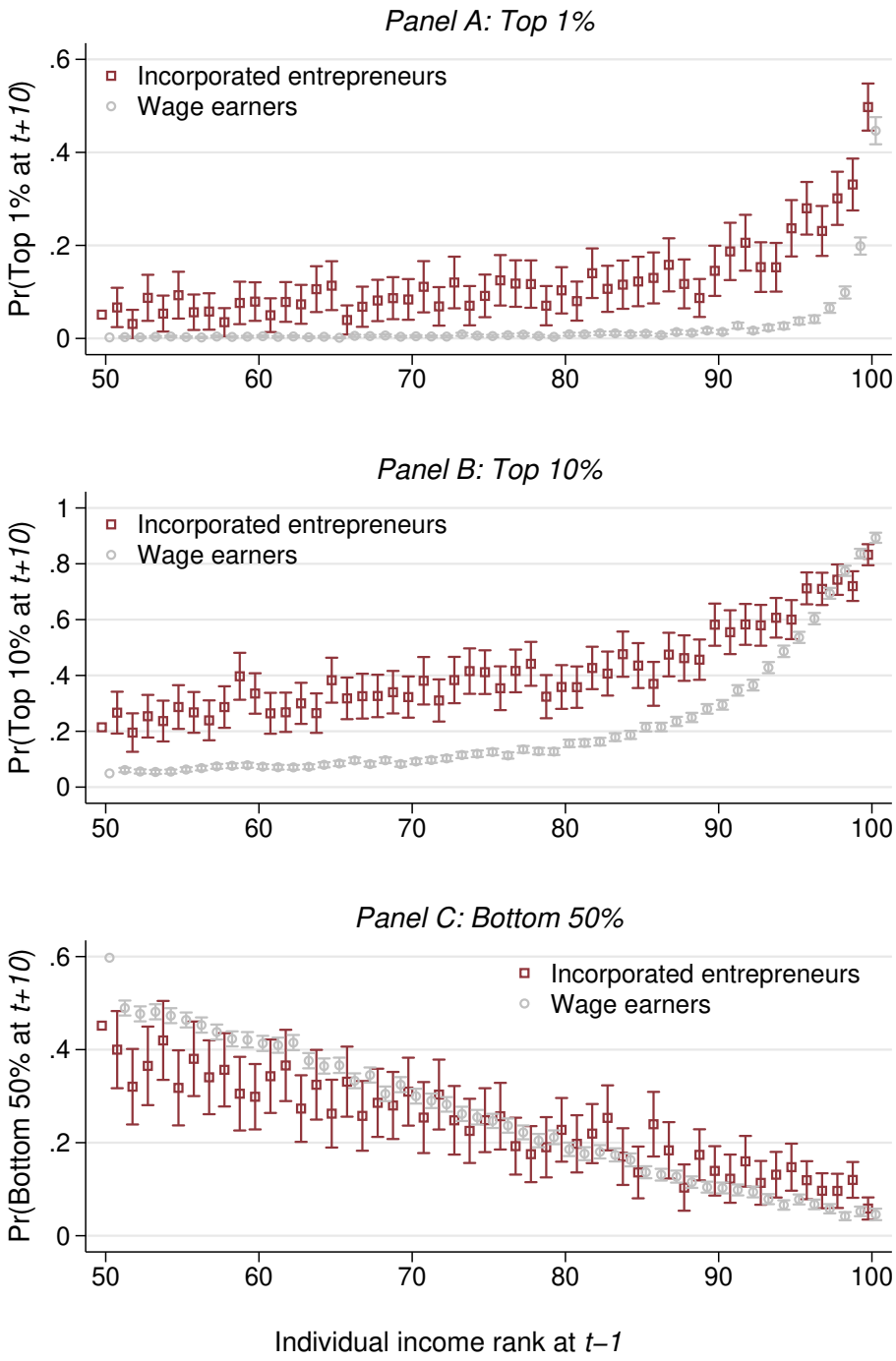


Notes: The figure presents the difference in log disposable income trajectories between incorporated entrepreneurs (IE) and wage earners (WE) before and after the business was established ($t = 0$), relative to one year before starting the first business $t - 1$ (denoted by zero in the figure). The figure shows the graphs for individual income trajectories for each quintile of the individual income rank. In each graph, the development is further split into five parental income rank quintiles. For wage earners, pseudo-starts at $t = 0$ are randomly drawn from a uniform distribution. Individual rank is defined as an average over the years $t - 5$ and $t - 1$ and parental income is observed when the parents are 45–50 years old. The figure shows that IE earn much more compared to WE after $t = 0$ in all income groups, but particularly among the group with both high individual and parental incomes.

Figure 10: Income ranks at $t-1$ and $t+10$ for new incorporated entrepreneurs and income ranks for all incorporated entrepreneurs 2006–2016

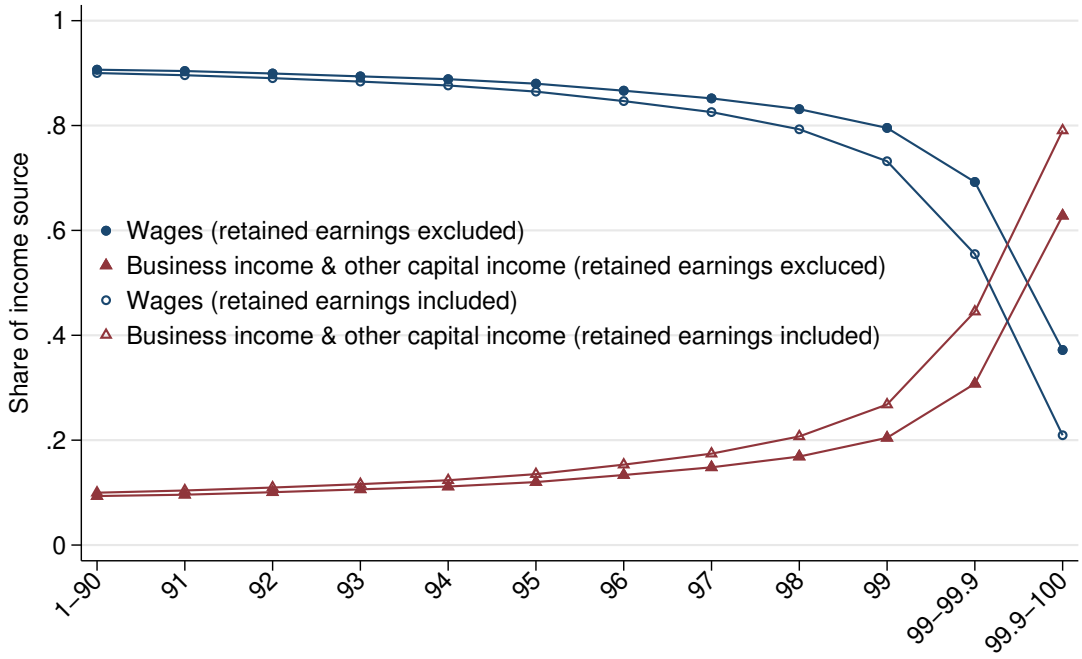
Notes: Figure presents the percentage shares of new incorporated business owners by their individual income ranks. The hollow diamonds show the shares one year before ($t-1$) and the hollow circles the shares 10 years ($t+10$) after establishing the business. The shares are plotted for those entrepreneurs that we observe in the data in both of these years. Individual income ranks are calculated from the distribution of disposable income in $t-1$ including all individuals in our baseline event study sample. The ranks of all incorporated entrepreneurs marked with red squares refer to the entire pool of them in 2006–2016 (not just new entrepreneurs), and ranks are calculated among the full population of Finnish individuals over 16 years of age. The figure shows that the share of entrepreneurs at the very top is large even before starting the business, but further increases during the first 10 years after the business was started. The shares at the top after 10 years resemble those in the overall distribution including all incorporated entrepreneurs.

Figure 11: Probability of reaching the top of the income distribution



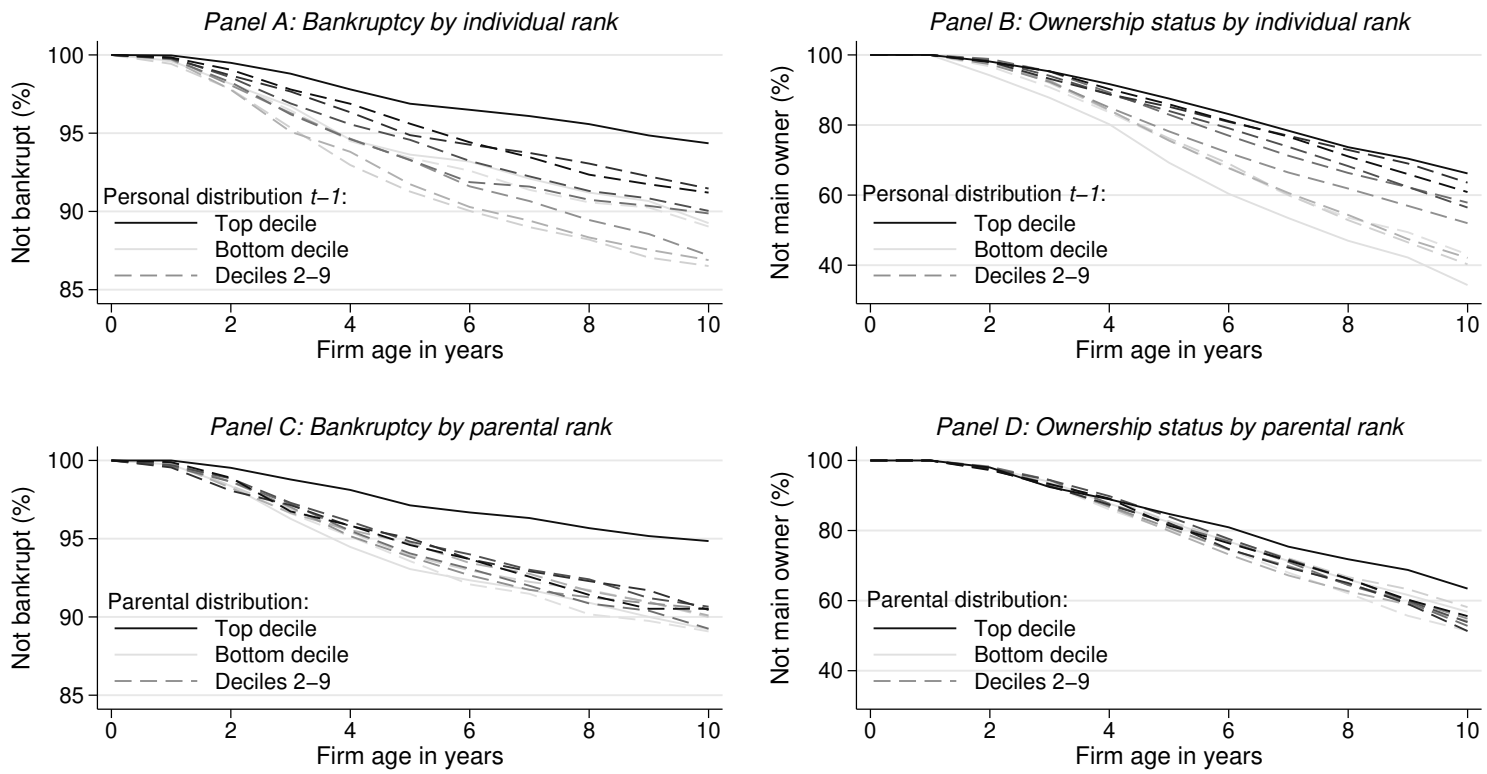
Notes: Figure plots the probabilities of locating in the top 1% (upper graph), top 10% (middle graph) and bottom 50% (bottom graph) of the income distribution 10 years after the business was established ($t + 10$), relative to the position in the distribution one year before ($t - 1$) for incorporated entrepreneurs (squares) and wage earners (circles) with 95% confidence intervals. For wage earners, pseudo-starts are drawn from a uniform distribution. The numbers at rank 50 represent the average probability among the bottom 50% in each graph. The figure shows that IEs have a higher probability of reaching the top compared to WEs throughout the income distribution, but there is no clear difference in the likelihood of locating in the bottom 50%.

Figure 12: Composition of market income sources by income percentiles



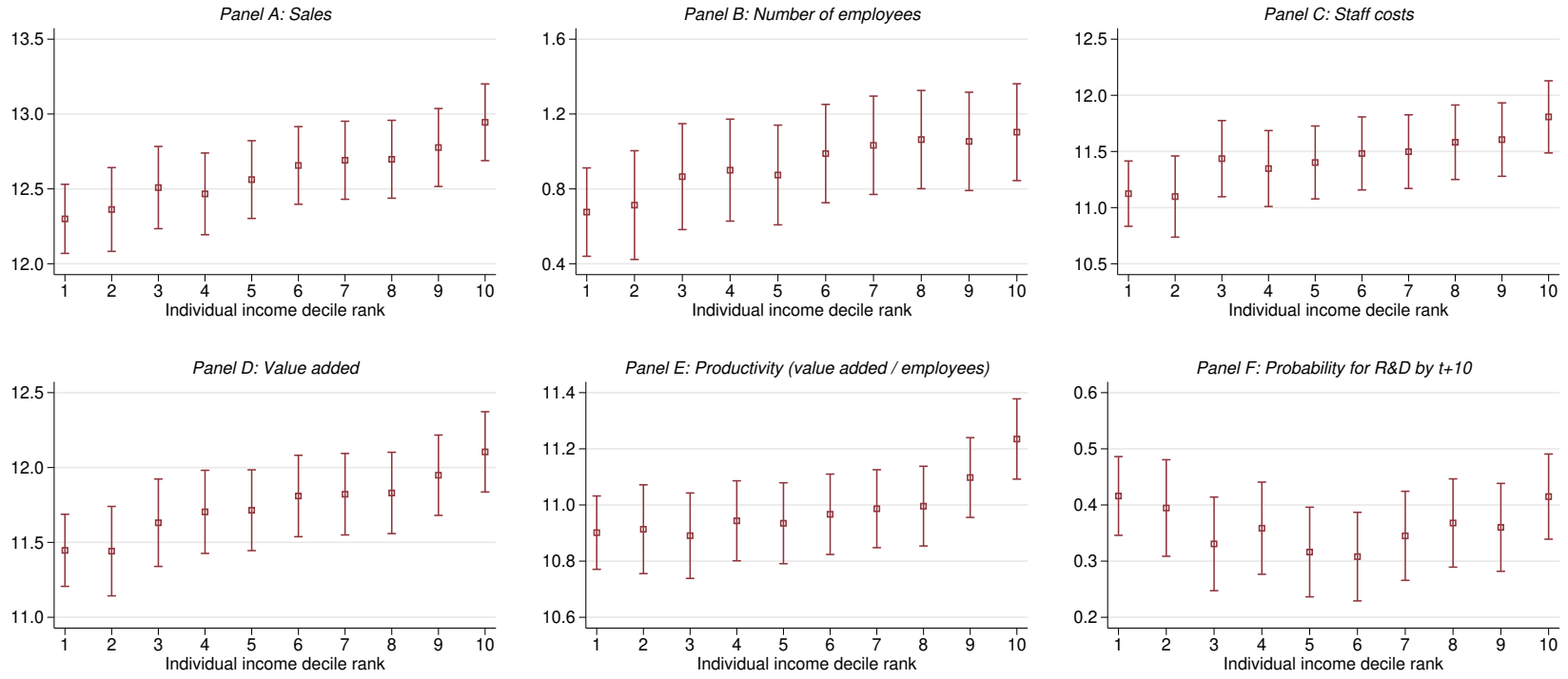
Notes: The figure plots the income composition between wages and business income + other capital income in the top 10% of the market income distribution for the full population of Finnish individuals over 16 years of age in 2006-2016. In addition, the figure includes the income shares when including in business income the annual retained earnings of the firms divided among each firm owner based on their ownership share of the firm. The detailed description of the variables is presented in Appendix D. The figure shows that the share of business income and other capital income increases sharply at the very top of the distribution in Finland, very similarly as in the US, as presented by Smith et al. (2019).

Figure 13: Survival rates of new incorporated businesses by individual and parental income ranks



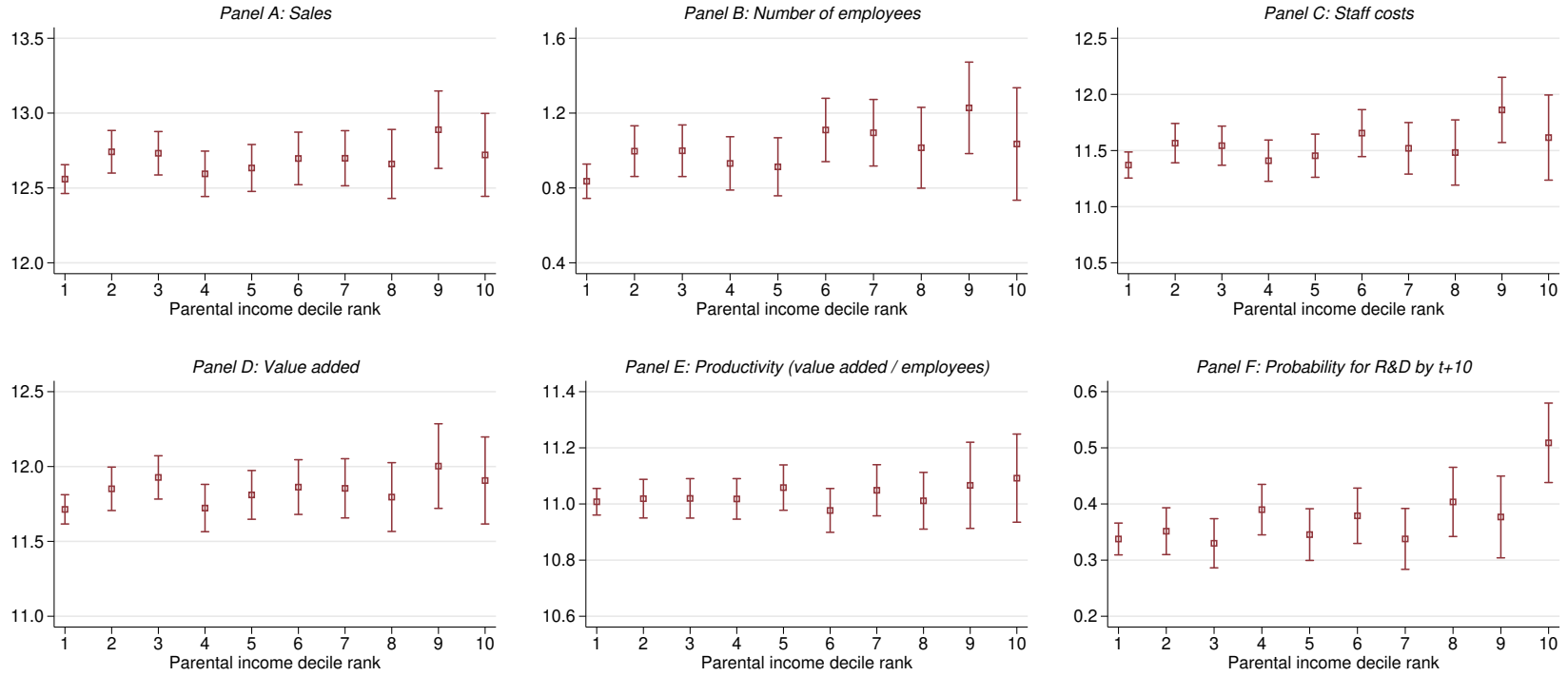
Notes: Panels A and B of the figure show the share of incorporated businesses that are not bankrupt in each firm age year (1–10) and by the main owner status of the business, indicating the share of new firms with the same initial main owner still holding main ownership in the firm, by the individual income ranks of the main owner. Panels C and D present similar graphs by parental income rank. The top income rank deciles are marked with a solid black line, and the bottom decile with a solid grey line. Deciles 2–9 are marked with dashed lines. The figure shows that the survival rates are higher using both measures among those with the highest parental income. The patterns are similar but not as pronounced for those with the highest individual income before starting a new business.

Figure 14: Firm-level outcomes by individual income ranks



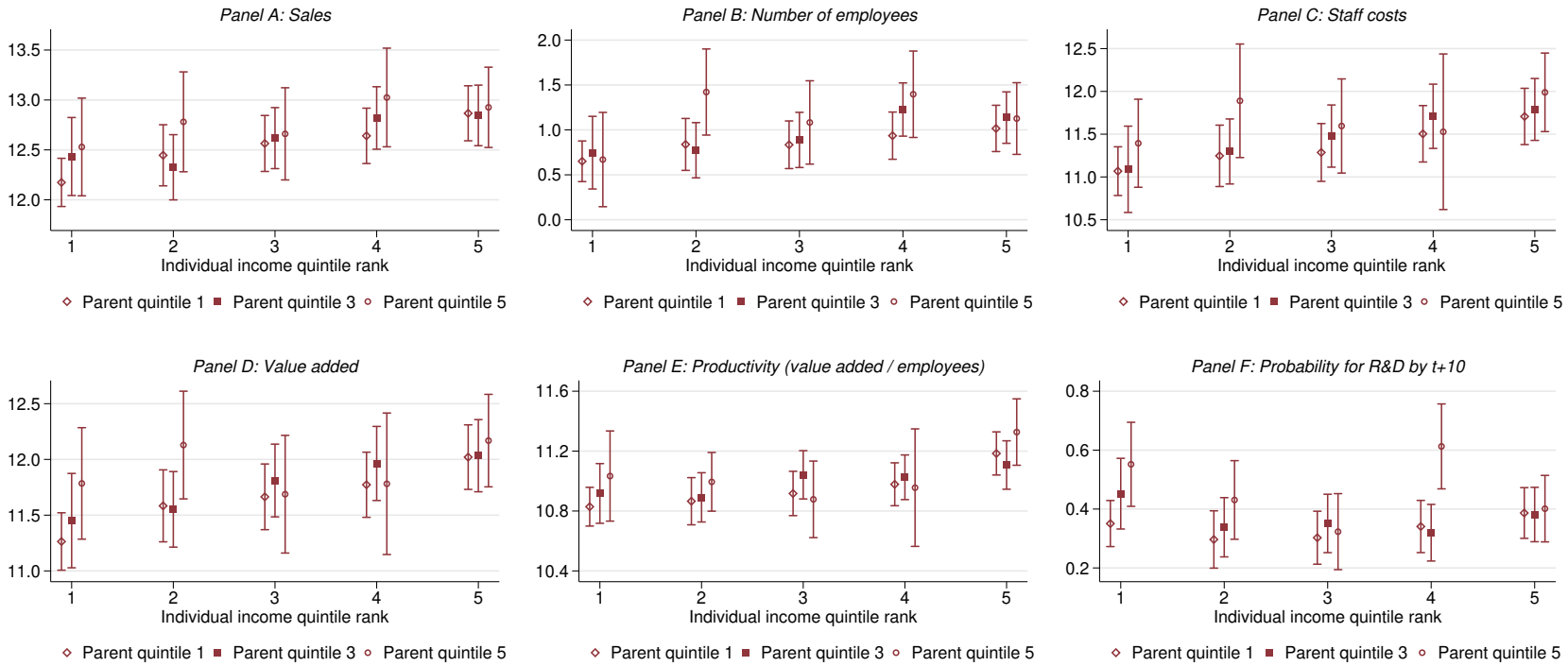
Notes: The figure shows firm-level outcomes with 95% confidence intervals by individual income ranks, defined as an average over the years $t-5$ and $t-1$. In Panels A-E, the outcomes are in log scale and are observed 10 years after the firm was established. Panel F presents the probability of positive R&D investments during the first ten years. The detailed description of the variables is presented in Appendix D. The figure shows that except for R&D investments, all of the firm outcomes increase with individual income rank measured before the firm was established.

Figure 15: Firm-level outcomes by parental income ranks



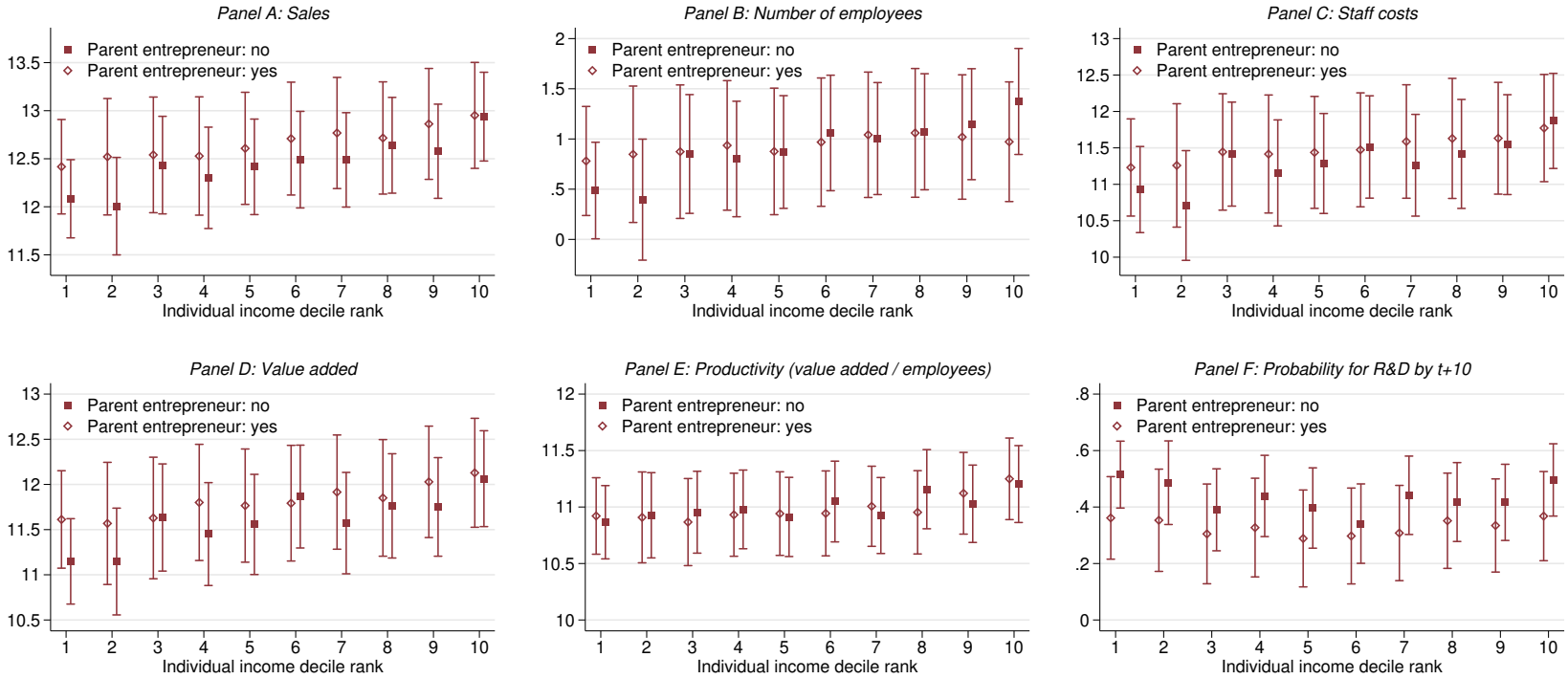
Notes: The figure shows firm-level outcomes with 95% confidence intervals by parental income ranks observed when the parents are 45–50 years old. In Panels A-E, the outcomes are in log scale and are observed 10 years after the firm was established. Panel F presents the probability of positive R&D investments during the first 10 years. The detailed description of the variables is presented in Appendix D. The figure shows that none of the firm outcomes are increasing by parental income rank, except for R&D investments, where the probability for these investments is higher for those with high parental income.

Figure 16: Firm-level outcomes by individual and parental income ranks



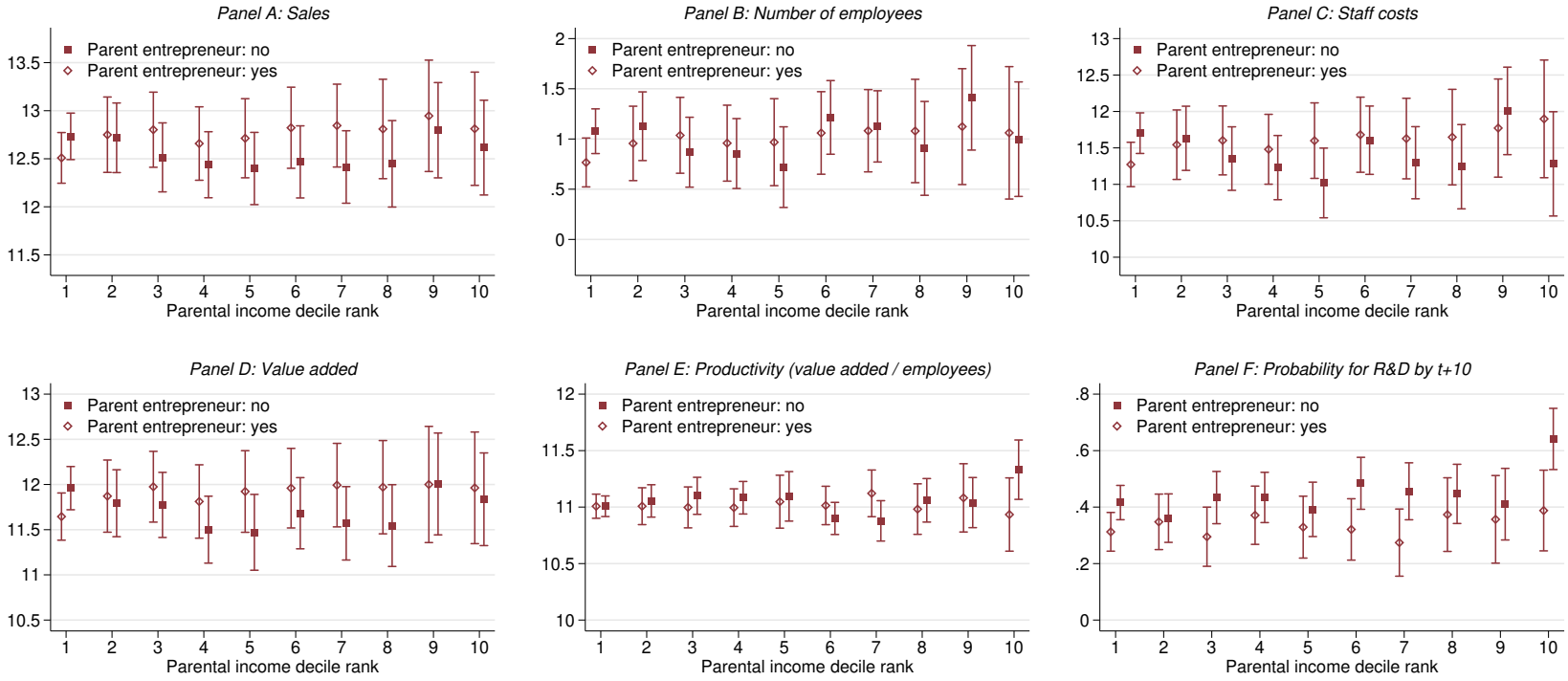
Notes: The figure shows firm-level outcomes with 95% confidence intervals by both individual and parental income distributions. The diamonds show the outcomes for those located in the first quintile of parental income distribution by different individual income quintiles (horizontal axis). The squares and circles show the outcomes for individuals in the third and fifth parental income quintiles, respectively. Individual income ranks are measured as an average over the years $t-5$ and $t-1$ from the full distribution including both wage earners and all business owners, and parental income ranks are observed when the parents are 45–50 years old. In Panels A-E, the outcomes are in log scale and observed 10 years after the firm was established. Panel F presents the probability of R&D investments during the first 10 years. The detailed description of the variables is presented in Appendix D. The figure shows a similar pattern as in Figures 14 and 15: firm-level outcomes increase with individual income but not by the parental income rank.

Figure 17: Firm-level outcomes by individual income ranks and parental entrepreneurship



Notes: The figure shows firm-level outcomes with 95% confidence intervals by individual income ranks and by parental entrepreneurship. The income ranks are measured as an average over the years $t-5$ and $t-1$ from the full distribution of new business owners and wage earners. Parents are classified as entrepreneurs if they are labeled as a business owner for at least 5 years according to the socio-economic status defined by Statistics Finland. In Panels A-E, the outcomes are in log scale and are observed 10 years after the firm was established. Panel F presents the probability for positive R&D investments during the first ten years. The detailed description of the variables is presented in Appendix D. The figure shows that firm-level outcomes do not vary by parental background in entrepreneurship in any of the individual income rank deciles, except for R&D investments, where the likelihood of these investments is slightly larger for those without entrepreneurial parents.

Figure 18: Firm-level outcomes by parental income ranks and parental entrepreneurship



Notes: The figure shows firm-level outcomes with 95% confidence intervals by parental income and parental entrepreneurship. The income ranks are calculated using the average of annual household income when the parents were 45–50 years old. Parents are classified as entrepreneurs if they are labeled as a business owner for at least 5 years according to the socio-economic status defined by Statistics Finland. In Panels A-E, the outcomes are in log scale and are observed 10 years after the firm was established. Panel F presents the probability for positive R&D investments during the first 10 years. The detailed description of the variables is presented in Appendix D. The figure shows that firm-level outcomes do not significantly vary by parental background in entrepreneurship in any of the parental income rank deciles, except for R&D investments, where the likelihood of these investments is slightly larger for those without entrepreneurial parents.

Appendix

A Additional Tables and Figures

Table A1: Number of new businesses and business owners in our sample, 1998–2014

	Number of firms or owners
Panel A: New incorporated firms	
All new firms	142,213
New firms with positive sales	129,703
New firms with positive sales and information about ownership structure	74,558
Panel B: Owners of new incorporated firms with positive sales in the first year of operation	
All owners	139,685
Main owners	95,109
Main owners with parental link in data	58,897
Main owners with parental link in data and no prior entrepreneurial experience (our baseline sample)	28,707
Panel C: New unincorporated firms	
All new firms	344,100
New firms with positive sales	316,063
New firms with positive sales and information about ownership structure	316,063
Panel D: Owners of new unincorporated firms with a positive sales in the first year of operation	
All owners	317,686
Main owners	316,950
Main owners who derive the majority of their annual earnings as business income	152,700
Main owners who derive the majority of their annual earnings as business income with parental link in data	102,140
Main owners who derive the majority of their annual earnings as business income with a parental link in data and no prior entrepreneurial experience (our baseline sample)	61,876

Notes: Table presents how the number of new businesses and their owners is shaped by our definitions and data availability. The bottom rows in Panels B and D correspond to our baseline samples of incorporated and unincorporated entrepreneurs.

Table A2: Descriptive statistics by income bin quintiles (in $t-1$)

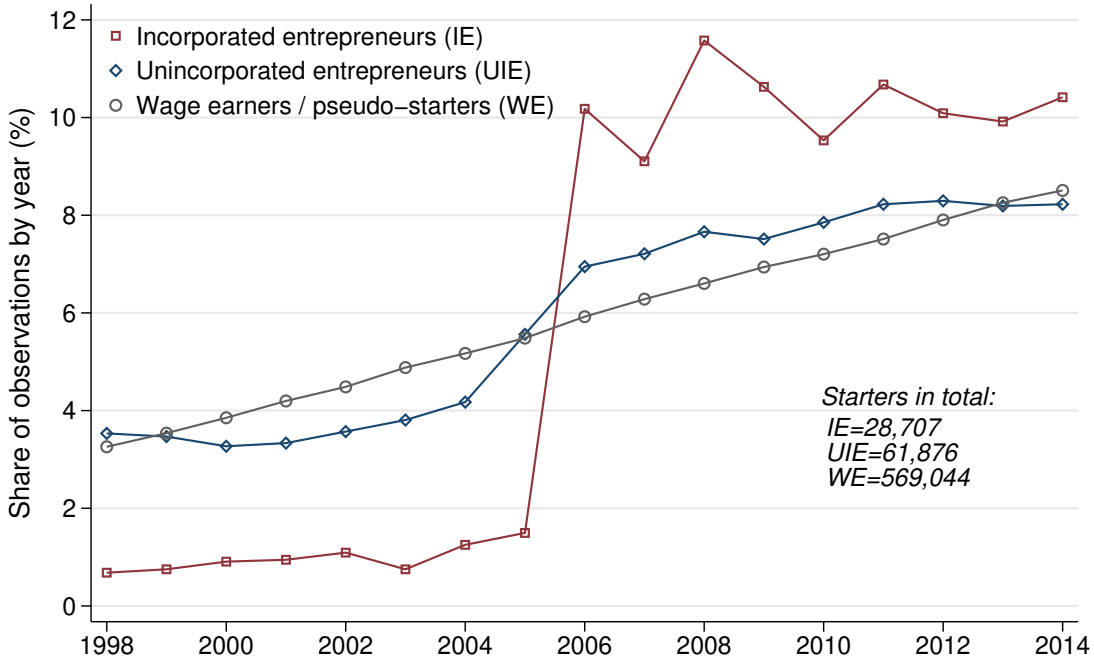
Income rank		Disposable	Market			Capital	Higher		
Individual	Parental	income	income	Age	Female	region	education	Unemployed	Student
Panel A: Incorporated entrepreneurs									
P1-P20	P1-P20	12,728	12,724	35	0.34	0.23	0.14	0.29	0.09
P1-P20	P21-P40	12,692	12,710	32	0.32	0.19	0.23	0.31	0.15
P1-P20	P41-P60	12,959	12,984	30	0.33	0.24	0.29	0.28	0.21
P1-P20	P61-P80	12,672	12,645	28	0.32	0.31	0.34	0.27	0.30
P1-P20	P81-P100	12,485	12,568	27	0.34	0.40	0.39	0.23	0.39
P21-P40	P1-P20	18,995	21,545	36	0.35	0.18	0.14	0.27	0.02
P21-P40	P21-P40	18,542	20,906	33	0.36	0.18	0.19	0.28	0.03
P21-P40	P41-P60	19,185	22,190	31	0.30	0.21	0.25	0.27	0.03
P21-P40	P61-P80	19,888	23,432	29	0.29	0.28	0.30	0.24	0.06
P21-P40	P81-P100	21,108	25,929	29	0.31	0.45	0.43	0.18	0.07
P41-P60	P1-P20	22,640	27,583	37	0.29	0.17	0.13	0.24	0.01
P41-P60	P21-P40	22,996	28,426	35	0.28	0.18	0.18	0.24	0.01
P41-P60	P41-P60	23,767	29,675	33	0.25	0.20	0.25	0.24	0.01
P41-P60	P61-P80	24,276	30,423	31	0.24	0.26	0.30	0.23	0.01
P41-P60	P81-P100	25,166	32,619	30	0.25	0.43	0.46	0.16	0.04
P61-P80	P1-P20	26,917	34,994	38	0.25	0.18	0.18	0.23	0.00
P61-P80	P21-P40	27,209	35,149	36	0.22	0.19	0.24	0.21	0.00
P61-P80	P41-P60	27,787	36,649	34	0.19	0.24	0.31	0.19	0.00
P61-P80	P61-P80	28,029	37,070	33	0.23	0.28	0.40	0.19	0.01
P61-P80	P81-P100	30,582	40,728	32	0.24	0.45	0.54	0.16	0.01
P81-P100	P1-P20	44,595	66,168	41	0.19	0.28	0.38	0.12	0.00
P81-P100	P21-P40	41,980	62,022	38	0.17	0.26	0.42	0.13	0.00
P81-P100	P41-P60	43,212	64,456	37	0.18	0.30	0.50	0.12	0.00
P81-P100	P61-P80	45,372	67,911	36	0.19	0.36	0.59	0.10	0.00
P81-P100	P81-P100	51,455	76,484	35	0.18	0.51	0.66	0.08	0.00
Panel B: Unincorporated entrepreneurs									
P1-P20	P1-P20	8,580	6,559	34	0.47	0.17	0.12	0.37	0.12
P1-P20	P21-P40	9,090	7,061	31	0.51	0.18	0.17	0.35	0.19
P1-P20	P41-P60	9,501	7,522	29	0.50	0.23	0.23	0.34	0.26
P1-P20	P61-P80	9,606	7,889	28	0.51	0.29	0.30	0.28	0.34
P1-P20	P81-P100	9,601	8,025	27	0.50	0.45	0.37	0.21	0.45
P21-P40	P1-P20	13,715	13,005	35	0.43	0.15	0.08	0.38	0.05
P21-P40	P21-P40	14,017	13,641	33	0.42	0.17	0.12	0.34	0.05
P21-P40	P41-P60	14,546	14,319	31	0.44	0.20	0.17	0.31	0.07
P21-P40	P61-P80	15,387	15,719	30	0.44	0.29	0.24	0.26	0.08
P21-P40	P81-P100	16,222	17,161	29	0.45	0.41	0.34	0.24	0.12
P41-P60	P1-P20	17,107	18,491	36	0.33	0.17	0.09	0.32	0.02
P41-P60	P21-P40	17,706	19,605	34	0.32	0.17	0.11	0.30	0.02
P41-P60	P41-P60	18,486	20,553	33	0.33	0.20	0.17	0.29	0.02
P41-P60	P61-P80	19,823	22,646	31	0.35	0.28	0.26	0.23	0.02
P41-P60	P81-P100	21,870	25,113	31	0.39	0.39	0.36	0.20	0.04
P61-P80	P1-P20	21,771	25,948	38	0.29	0.18	0.12	0.25	0.01
P61-P80	P21-P40	23,000	27,745	36	0.27	0.19	0.16	0.25	0.01
P61-P80	P41-P60	22,997	27,711	34	0.29	0.27	0.22	0.23	0.01
P61-P80	P61-P80	24,235	29,187	33	0.31	0.32	0.35	0.20	0.01
P61-P80	P81-P100	25,515	30,733	32	0.40	0.45	0.42	0.21	0.02
P81-P100	P1-P20	34,296	47,360	40	0.33	0.23	0.35	0.16	0.00
P81-P100	P21-P40	35,732	49,651	38	0.29	0.23	0.38	0.15	0.00
P81-P100	P41-P60	37,810	53,431	37	0.33	0.32	0.47	0.12	0.00
P81-P100	P61-P80	40,326	57,289	35	0.43	0.36	0.60	0.11	0.00
P81-P100	P81-P100	45,497	62,803	35	0.44	0.46	0.72	0.08	0.01

Table A2: Descriptive statistics by income bins (continued)

Income rank		Disposable	Market			Capital	Higher		
Individual	Parental	income	income	Age	Female	region	education	Unemployed	Student
Panel C: Wage earners									
P1-P20	P1-P20	10,476	8,081	34	0.57	0.16	0.13	0.42	0.17
P1-P20	P21-P40	11,014	9,246	30	0.60	0.18	0.21	0.34	0.26
P1-P20	P41-P60	11,286	9,949	28	0.59	0.22	0.27	0.27	0.35
P1-P20	P61-P80	11,574	10,609	27	0.59	0.29	0.33	0.20	0.44
P1-P20	P81-P100	11,392	10,607	27	0.56	0.43	0.35	0.13	0.54
P21-P40	P1-P20	16,100	16,894	36	0.61	0.16	0.12	0.28	0.04
P21-P40	P21-P40	16,668	17,952	32	0.61	0.18	0.20	0.24	0.06
P21-P40	P41-P60	17,201	19,028	30	0.60	0.22	0.29	0.19	0.08
P21-P40	P61-P80	18,036	20,548	29	0.59	0.30	0.38	0.15	0.10
P21-P40	P81-P100	18,753	21,957	28	0.58	0.43	0.49	0.11	0.12
P41-P60	P1-P20	19,517	22,925	38	0.57	0.17	0.13	0.17	0.01
P41-P60	P21-P40	20,059	23,864	35	0.54	0.19	0.20	0.14	0.01
P41-P60	P41-P60	20,827	25,051	33	0.53	0.24	0.29	0.12	0.02
P41-P60	P61-P80	21,897	26,689	31	0.52	0.31	0.42	0.10	0.02
P41-P60	P81-P100	22,991	28,777	30	0.56	0.46	0.54	0.07	0.02
P61-P80	P1-P20	23,114	29,268	39	0.48	0.19	0.18	0.10	0.00
P61-P80	P21-P40	23,773	30,399	36	0.45	0.20	0.24	0.09	0.00
P61-P80	P41-P60	24,795	32,180	34	0.43	0.25	0.34	0.08	0.01
P61-P80	P61-P80	26,016	33,860	33	0.45	0.33	0.48	0.07	0.01
P61-P80	P81-P100	27,312	35,949	33	0.49	0.48	0.62	0.06	0.01
P81-P100	P1-P20	31,898	45,942	40	0.34	0.25	0.38	0.05	0.00
P81-P100	P21-P40	32,646	47,275	38	0.30	0.26	0.42	0.04	0.00
P81-P100	P41-P60	34,250	50,109	37	0.30	0.32	0.52	0.04	0.00
P81-P100	P61-P80	36,603	53,764	36	0.34	0.42	0.64	0.04	0.00
P81-P100	P81-P100	40,482	59,729	35	0.38	0.56	0.74	0.03	0.00

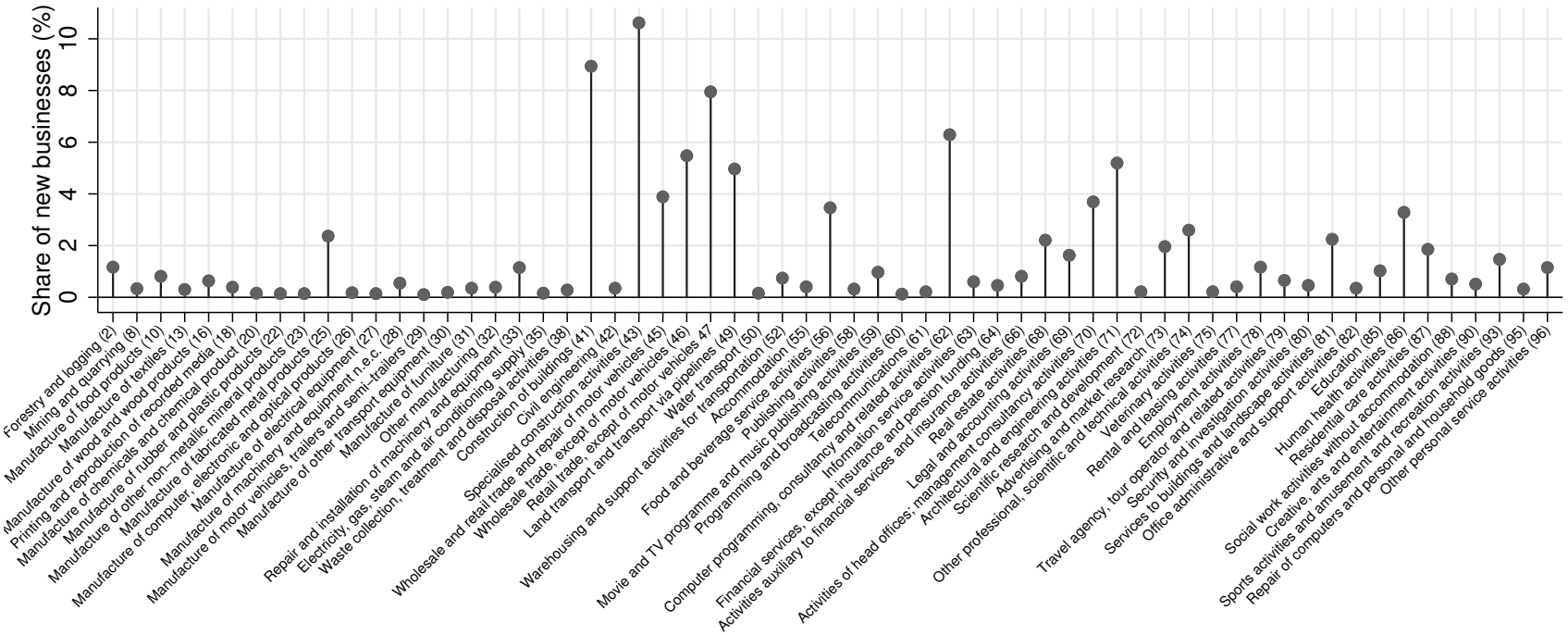
Notes: Table presents descriptive statistics for incorporated and unincorporated business owners and wage earners by income bin quintiles in year $t-1$, one year before the business was established and one year before the randomly allocated pseudo-starts for wage earners for the years 1997–2013. Individual rank is defined as an average over the years $t-5$ and $t-1$ while parental income is observed when the parents are 45–50 years old. Individuals are then grouped into five individual and five parental income quintiles, dividing the data into 25 bins altogether. Capital city region includes Helsinki, Espoo, Vantaa and Kauniainen. Higher education refers to a bachelor's degree or higher. Individuals are defined as unemployed or students, if they received any unemployment or student benefits, respectively, in year $t-1$. Detailed definitions for each variable are available in Appendix D.

Figure A1: New entrepreneurs over the sample years



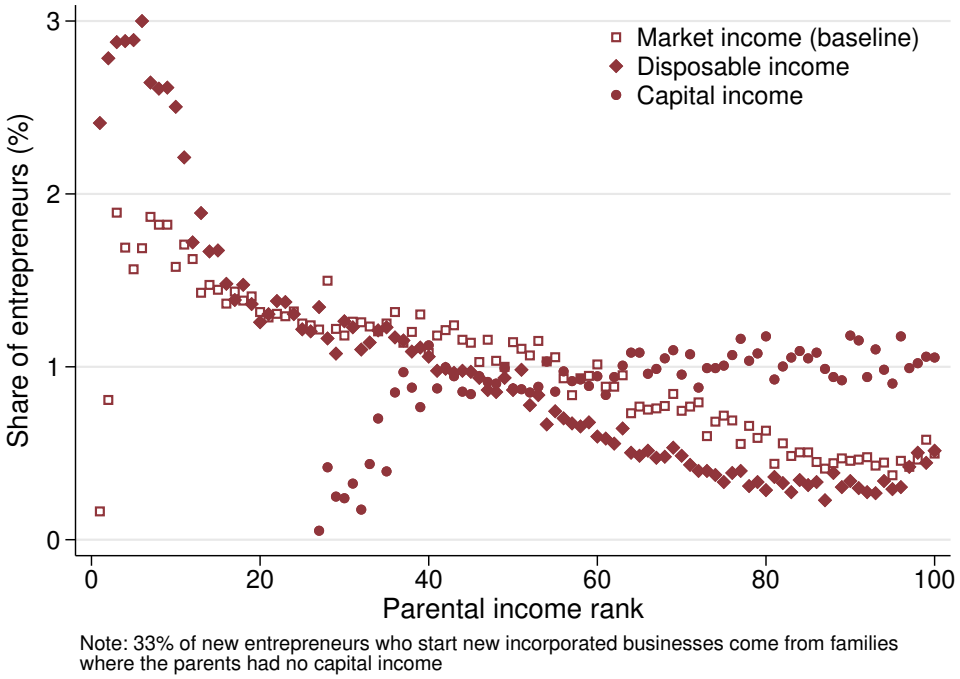
Notes: Figure presents the shares of new owners of incorporated and unincorporated new firms and the shares of wage earners (pseudo-starters) included in our baseline sample by year. The discontinuous jump in the number of new incorporated entrepreneurs stems from the change in the data source for defining new business owners in 2006 (see Section 2.2 for more details). The jump in the number of new owners does not significantly affect our results, as discussed in more detail in Appendix B.

Figure A2: Industry composition of newly established incorporated businesses



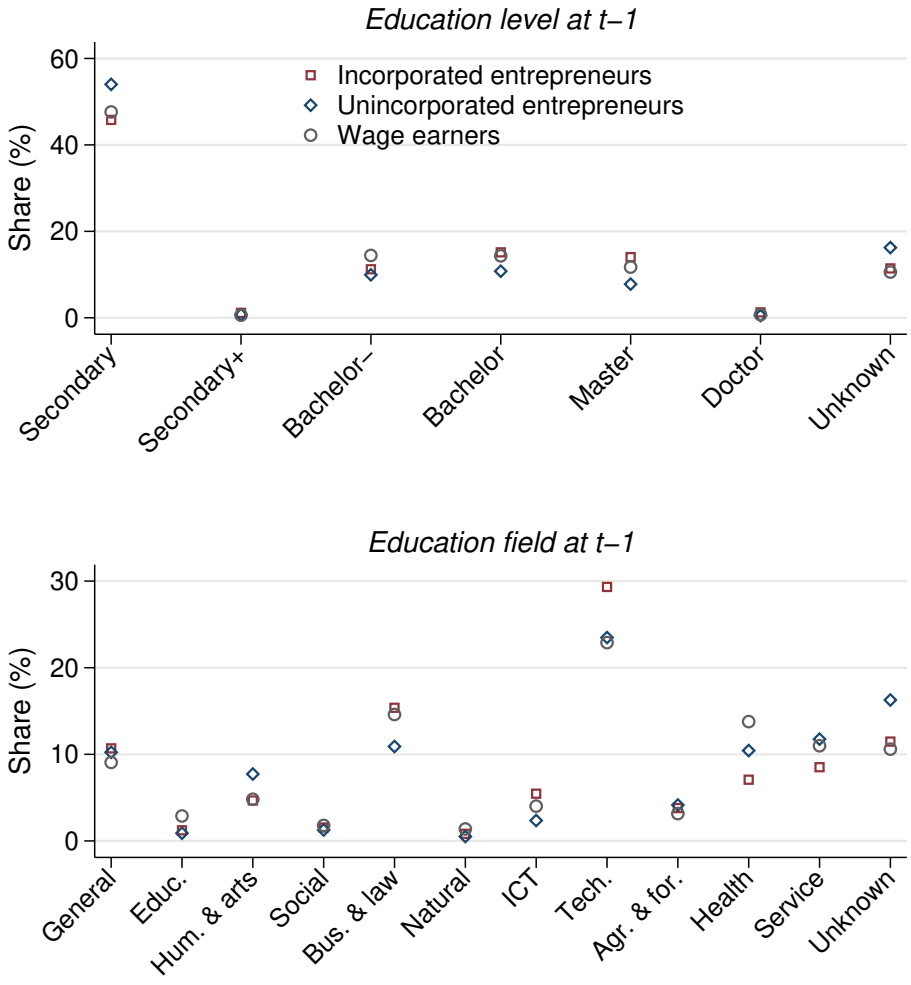
Notes: The figure shows the industry shares among the newly established incorporated businesses in our baseline sample using the TOL 2008 classification by Statistics Finland, which is based on the EU's classification of economic activities (NACE).

Figure A3: Share of new entrepreneurs by parental income ranks using different income concepts



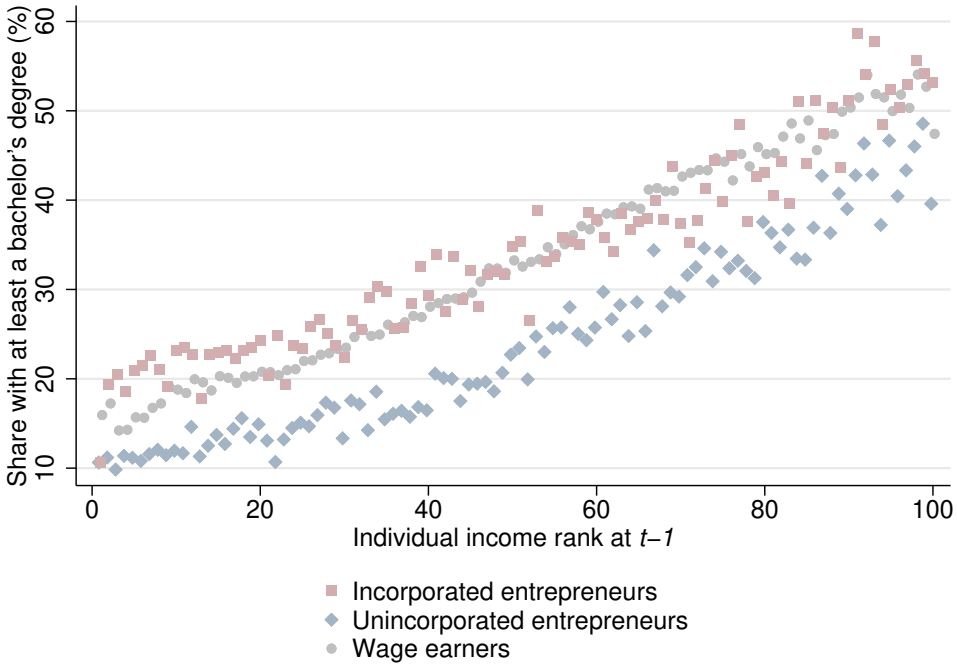
Notes: Figure presents the percentage shares of new incorporated business owners by their parental income ranks using three different income concepts. As our baseline measure (squares), parental income is calculated as average annual household market income when the parents were 45–50 years old. For a subset of parents who are 45–50 years old after 1995, we can define parental ranks using disposable income (diamonds) and capital income (circles). The figures are very similar for market income and disposable income, except that for disposable income there is a small increase in the share of new entrepreneurs at the very top. For capital income, the share of new entrepreneurs remains flat from the 40th rank percentile onward. Approximately 25% of the parents do not have positive capital income in the data.

Figure A4: Share of new entrepreneurs by education



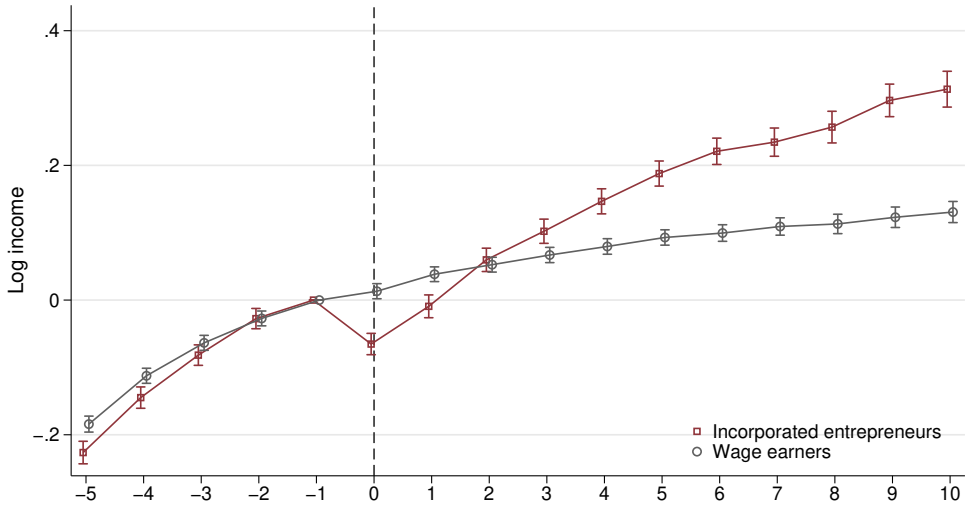
Notes: Figure presents the percentage shares of new incorporated (red squares) and unincorporated (blue diamonds) business owners by their education one year before establishing the business ($t-1$). The share of wage earners (grey dots) indicates the share of pseudo-starters. Education level and field correspond to ISCED 2011 classification (number in parenthesis). Secondary stands for lower secondary education (2), Secondary+ for upper secondary education (3) or post-secondary non-tertiary education (4), Bachelor- for short-cycle tertiary education (5), Bachelor for bachelor's degree or equivalent level (6), Master for master's degree or equivalent level (7), Doctor for doctoral degree or equivalent level (8), and Unknown for those that are not elsewhere classified (9). For the education fields, General stands for generic programmes and qualifications (00), Educ. for education (01), Hum. & arts for humanities and arts (02), Social for social sciences, journalism and information (03), Bus. & law for business, administration and law (04), Natural for natural sciences, mathematics and statistics (05), ICT for information and communication technologies (06), Tech. for engineering, manufacturing and construction (07), Agr. & for. for agriculture, forestry, fisheries and veterinary (08), Health for health and welfare (09), Service for services (10), and Unknown for unknown categories (99). The figure shows that there are no notable differences in education levels between the groups. The shares in different fields are also similar, but new incorporated entrepreneurs are more likely to have engineering, manufacturing or construction as their field of education compared to the other groups.

Figure A5: Share of highly-educated individuals by individual income ranks



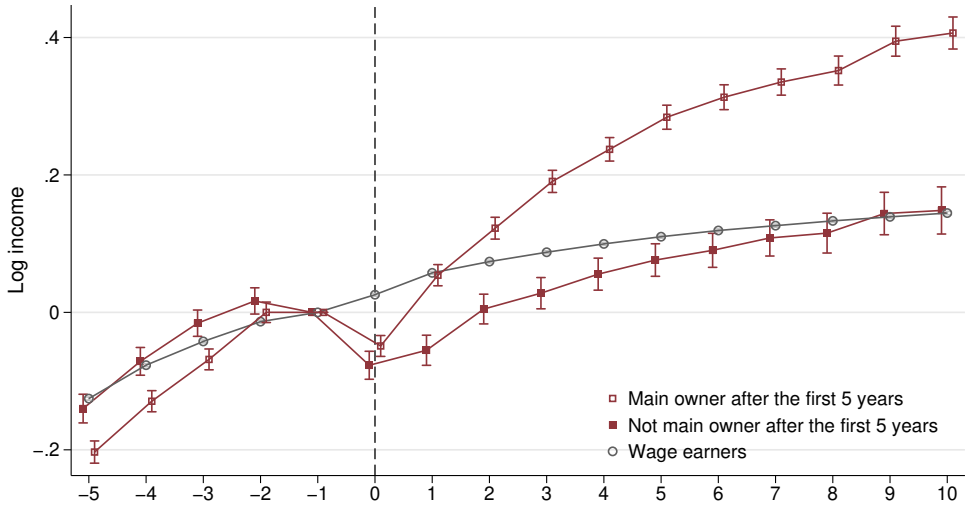
Notes: Figure presents the percentage shares of incorporated and unincorporated entrepreneurs who have at least a bachelor's degree by their individual income ranks one year before establishing the business $(t-1)$, and for wage earners one year before the randomly allocated pseudo-start years. The figure shows a similarly increasing pattern of education by individual income rank in all of the groups.

Figure A6: Disposable income trajectories, coarsened exact matching (CEM)



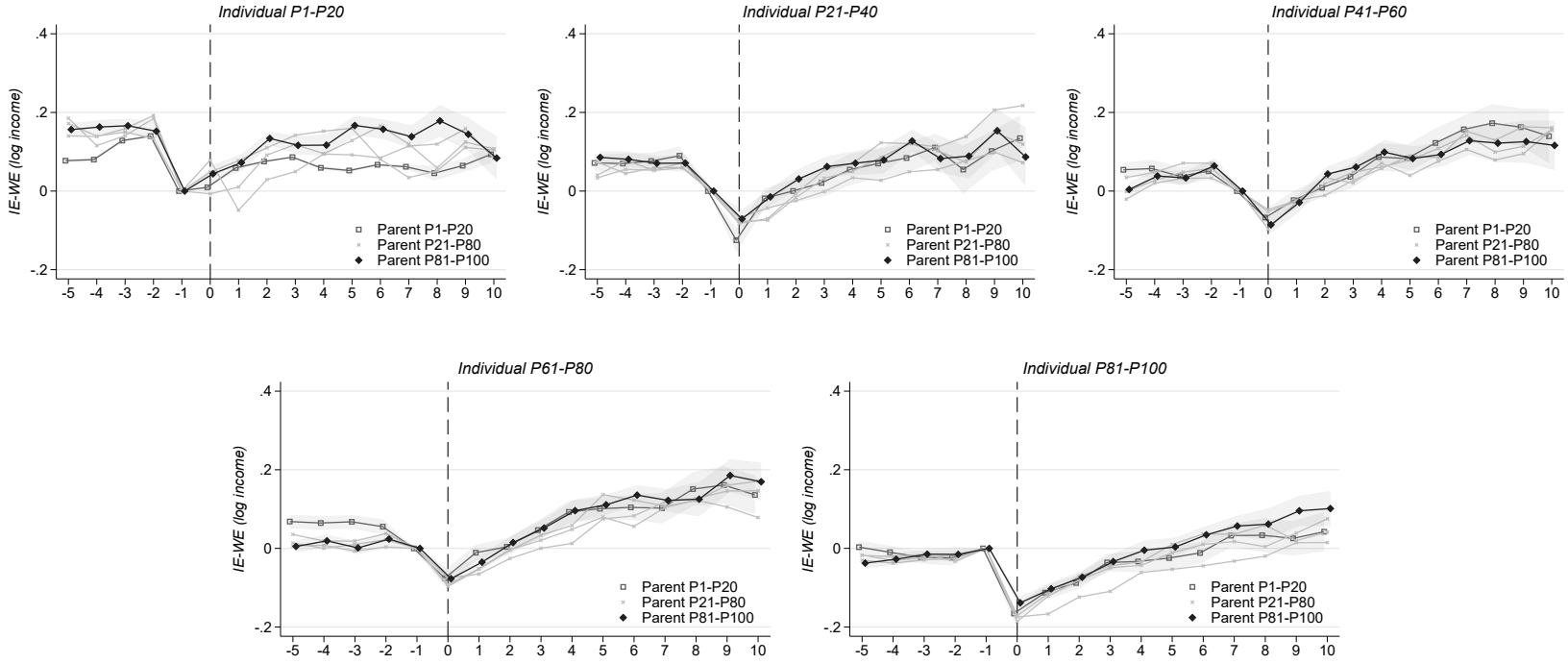
Notes: The figure presents the disposable income trajectories of incorporated business owners (red squares) and wage earners (gray circles) with 95% standard errors before and after the business was established ($t = 0$), relative to the year $t - 1$ (denoted by zero in the figure). For wage earners, pseudo-start years at $t = 0$ are randomly drawn from a uniform distribution. The two groups are matched using coarsened exact matching (CEM). We use the average decile income rank in $t-3$ and $t-1$, the change in rank between $t-3$ and $t-1$, parental income rank, an indicator variable for one or two parents being an entrepreneur, average share of labor earnings relative to total income in $t-3$ and $t-1$, and age, gender, place of residence and number of children in $t-1$ in the matching procedure. The CEM procedure removes most of the difference in the trends before $t = 0$ between the groups, and shows that even conditional on matching, incorporated entrepreneurs experience significantly higher income growth after establishing their first business.

Figure A7: Income trajectories by firm survival five years after establishment



Notes: The figure presents the unconditional disposable income trajectories of incorporated business owners separately for entrepreneurs who are still the main owners of their newly established business five years afters the business was established (hollow red squares), and those who are not (filled red squares). The trajectories are portrayed before and after the business was established ($t = 0$), relative to the year $t - 1$ (denoted by zero in the figure). For wage earners (gray circles), pseudo-start years are randomly drawn from a uniform distribution.

Figure A8: Difference in disposable income trajectories between incorporated entrepreneurs and wage earners by individual and parental income with Mincer controls



Notes: The figure presents the difference in disposable income trajectories between incorporated business owners and wage earners before and after the business was established ($t = 0$), relative to the year $t-1$ (denoted by zero in the figure) by individual and parental income rank quintiles similarly as in Figure 9 in the main text but with Mincer-controls. Individual rank is defined as an average over the years $t-5$ and $t-1$ while parental income is observed when the parents are 45–50 years old. The Mincer-style controls include all interactions of age, sex and education (primary education, secondary degree or tertiary degree) and the interactions using age squared and age cubed.

B Robustness Checks and Sensitivity Analysis

B.1 Market Income and Retained Earnings

Focusing on market income instead of disposable income (our baseline measure) does not significantly affect the qualitative aspects of our analysis. Figure B1 below paints a very similar picture as Figure 7 did for disposable income in the main text. Market income includes the following income sources: wages, entrepreneurial income and capital income, which is the sum of dividends, interest income, private pensions, rental income on properties, taxable capital gains and other capital income.

In addition to market income, Figure B1 includes trajectories for market income plus the earnings retained in the firm. Defining an exact boundary between the firm and the owner's incomes can be very challenging (see e.g. Kopczuk and Zwick 2020). In addition to income withdrawn from the firm as wages, dividends and capital gains, the owners of privately held corporations can retain earnings in their firm. These earnings do not show up as income for the owners in the administrative data in the year that they are accrued.[14]

To study retained earnings, we follow a similar type of approach as in Alstadsæter et al. (2016) and allocate them to each owner of an incorporated firm based on their ownership share of the firm, and include this income on top of the market income observed in the administrative data.[15]

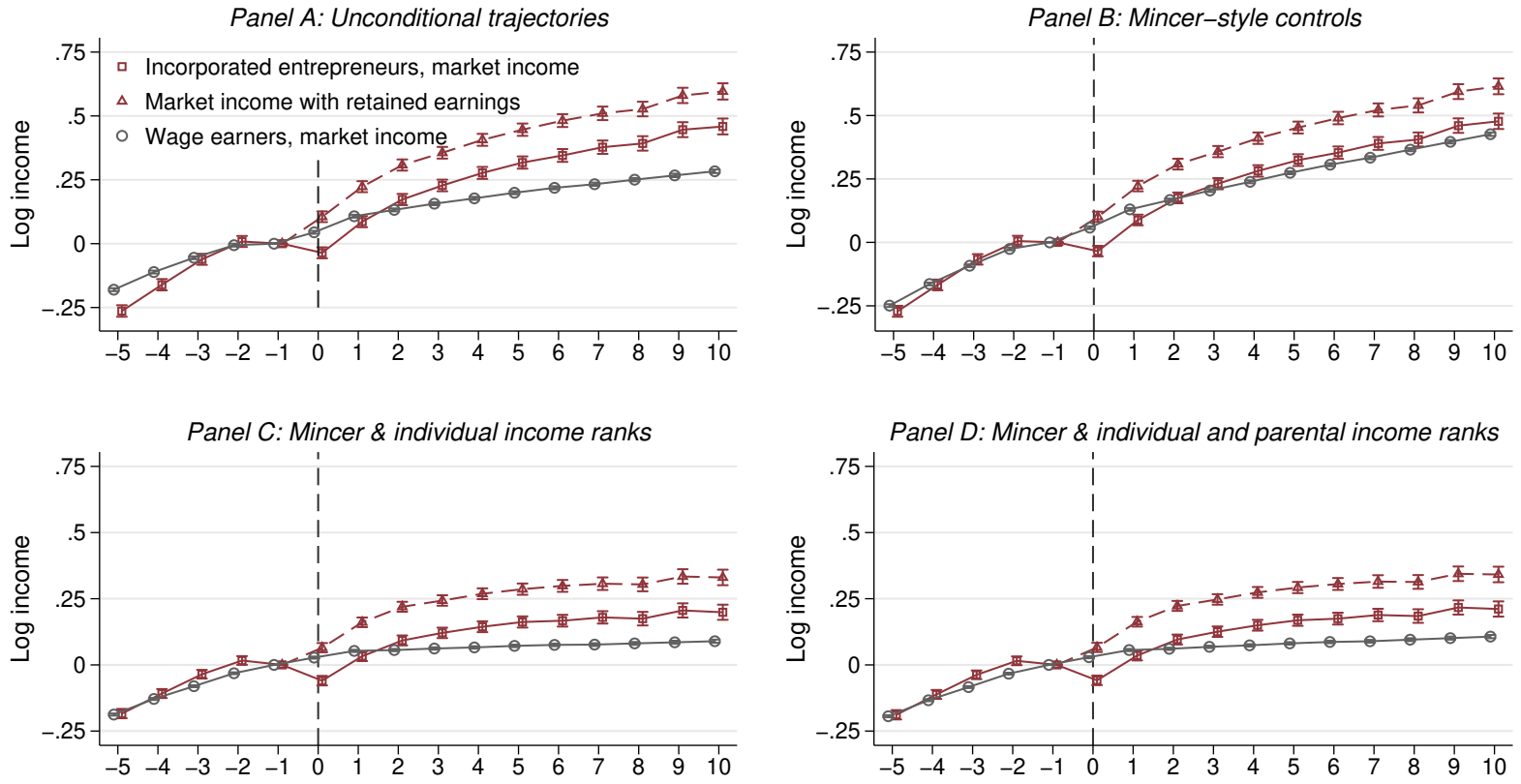
Figure B1 shows the income development for IEs and WEs when accounting for earnings retained in the firm.[16] Including retained earnings appears to increase the income levels of IEs by an almost constant annual share of 15%, starting right from the first year after the business was established.

[14]For example, income retained in the firm might show up in the reported individual income of the owner only after a long time period in a very lumpy fashion, making it difficult to derive a consistent income trajectory for entrepreneurs. Second, some of the income retained in the firm can be used for the owners' private consumption without ever showing up in the individual-level administrative data.

[15]Similarly as in the Norwegian institutional setting studied by Alstadsæter et al. (2016), the Finnish dividend tax system incentivizes owners to accumulate earnings in their firms. This is due to the fact that firm-level net assets (assets - debts) affect the owner-level dividend income tax rates such that a larger net assets position reduces the dividend tax rate of the owner, as described in more detail in Appendix C. Therefore, it is particularly important to take retained earnings into account as a potential income source for incorporated business owners in the Finnish context.

[16]As retained earnings are not relevant for UIEs in the Finnish context (all business income is taxed each year at the individual level for UIEs), Figure B1 only concentrates on comparing the income trajectories of IEs and WEs over time.

Figure B1: Average market income trajectories when accounting for retained earnings



Notes: Figure presents the log market income trajectories (red squares) and market income + retained earnings trajectories (red triangles) for incorporated entrepreneurs, and market income trajectories for wage earners (gray circles) before and after the business was established ($t = 0$), relative to the year $t - 1$ (denoted by zero in the figure) with 95% standard errors. The figure shows the unconditional trajectories (Panel A), trajectories with Mincer-controls (Panel B), Mincer and individual income rank controls (Panel C), and trajectories with Mincer and individual and parental income rank controls (Panel D). Retained earnings include the share of annual retained earnings of the firm allocated for each owner using their ownership share of the firm. For wage earners, pseudo-starts are randomly drawn from a uniform distribution. The Mincer-style controls include all interactions of age, sex and education (primary education, secondary degree or tertiary degree) and the interactions using age squared and age cubed. Individual income ranks are measured as an average over the years $t - 5$ and $t - 1$ from the full distribution including both wage earners and business owners in our baseline sample. Parental income ranks are calculated using the average of annual household income when the parents were 45–50 years old.

B.2 Other Robustness Checks and Heterogeneity

Table B1 below summarizes the average increase in disposable income among incorporated entrepreneurs from one year before becoming an entrepreneur to 10 years after using different subsamples of entrepreneurs and different estimation samples.

First, we find that the gains from entrepreneurship are somewhat larger among those who were wage earners right before establishing their first incorporated business than among those who were unincorporated entrepreneurs. Second, the table further confirms the result that parental entrepreneurship is not driving the differences in gains between entrepreneurs and wage earners. Third, the table shows that those who remained entrepreneurs for at least five years gained on average roughly 10% more than those who did not.

The bottom three elements in Table B1 illustrate the robustness of the results in terms of the estimation sample. As discussed in Section 2, we rely on data from the Finnish Tax Administration to define the owners of incorporated firms using information on dividend income received from their firm for the years 1998–2005 (Tax Administration data). For the years 2006–2014 we use data from Statistics Finland including all owners of incorporated firms (Statistics Finland data).

First, Table B1 shows that the observed gains are larger for the firms that started in 1998–2005 compared to 2006–2014. This could imply that the differences in the estimation sample of business owners over time is driving this difference. However, the first time period also coincides with rapid aggregate economic growth in the Finnish economy. In 1998–2005, the annual average GPD growth was 3.6%, while in 2006–2014 it was only 0.6%. To distinguish between the differences in the overall economic development and the estimation sample, we estimate the income gains using the definition based on the Tax Administration data for the years 2006–2014 in the bottom row of Table B1. Reassuringly, these results are nearly identical to those derived using the full owner sample from the Patent and Registration Office data within the same time period, illustrating that the differences in income gains over time are not driven by differences in the estimation samples.

Figure B2 shows the disposable income trajectories separately for our baseline sample (Panel A) and for the sample including individuals who we observe each year from $t-1$ to $t+10$ (Panel B), and from $t-5$ to $t+10$ (Panel C). The figure shows that the gains from entrepreneurship are higher in the balanced panels compared to our baseline non-balanced sample. First, keeping only individuals who are in our sample throughout the 10 first years overweights younger individuals, who tend to have steeper income trajectories than older individuals. Second, the balanced panel has a higher share of individuals starting their businesses during the economic upswing of the early and mid-2000s, which is associated with larger increases in income.

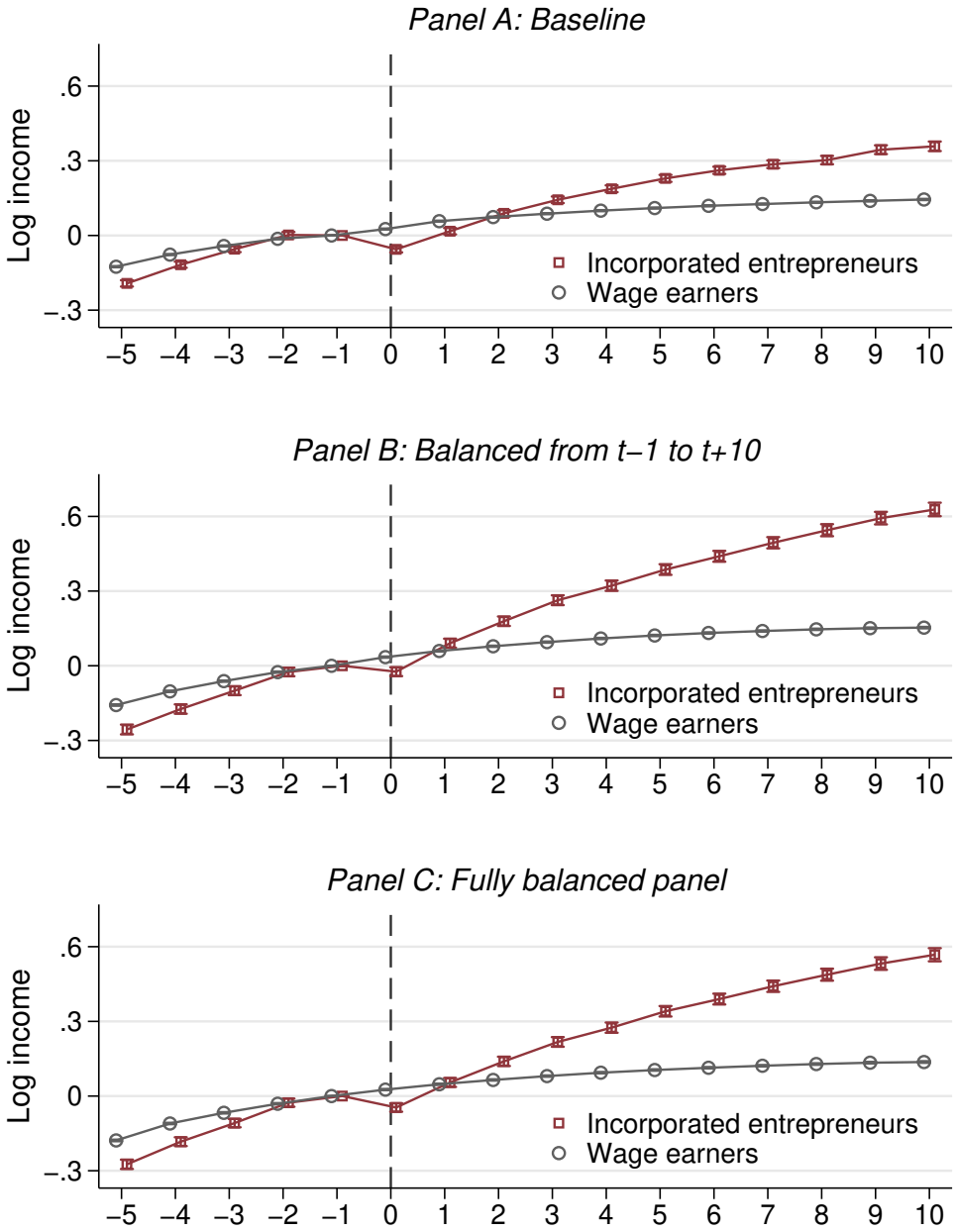
Finally, we illustrate how the gains from entrepreneurship differ between the industries of newly established firms in Figure B3. There is significant overall variation in average gains over different industries, but we do not observe any consistent differences between the more traditional industries such as manufacturing or transportation, and industries characterized by highly specialized human capital skills, such as consultancy or health services.

Table B1: Sensitivity of the gains from entrepreneurship to the estimation sample

		Mean disposable income (standard error)		
Sample	Observations	t-1	t+10	% Increase
Baseline sample	13,467	28,390	42,461	29.1%
		(293)	(674)	(0.8%)
Switchers (UIE-IE) dropped	11,538	27,933	43,384	33.0%
		(339)	(789)	(0.9%)
Parent entrepreneur: yes	7,697	28,000	42,453	29.0%
		(335)	(670)	(1.2%)
Parent entrepreneur: no	5,770	28,925	42,473	29.3%
		(520)	(1,310)	(1.2%)
Main owner after 5 years: yes	8,873	28,248	44,585	32.5%
		(334)	(915)	(1.0%)
Main owner after 5 years: no	4,594	28,667	38,279	22.5%
		(567)	(871)	(1.6%)
Firm starts 1998-2005 (Tax Admin. data)	4,932	29,204	47,591	37.2%
		(592)	(969)	(1.4%)
Firm starts 2006-2014 (Statistics Finland data)	8,535	27,911	39,444	24.4%
		(307)	(905)	(1.0%)
Firm starts 2006-2014 (Tax Admin. data)	1,654	32,875	45,517	24.9%
		(632)	(1,131)	(2.0%)

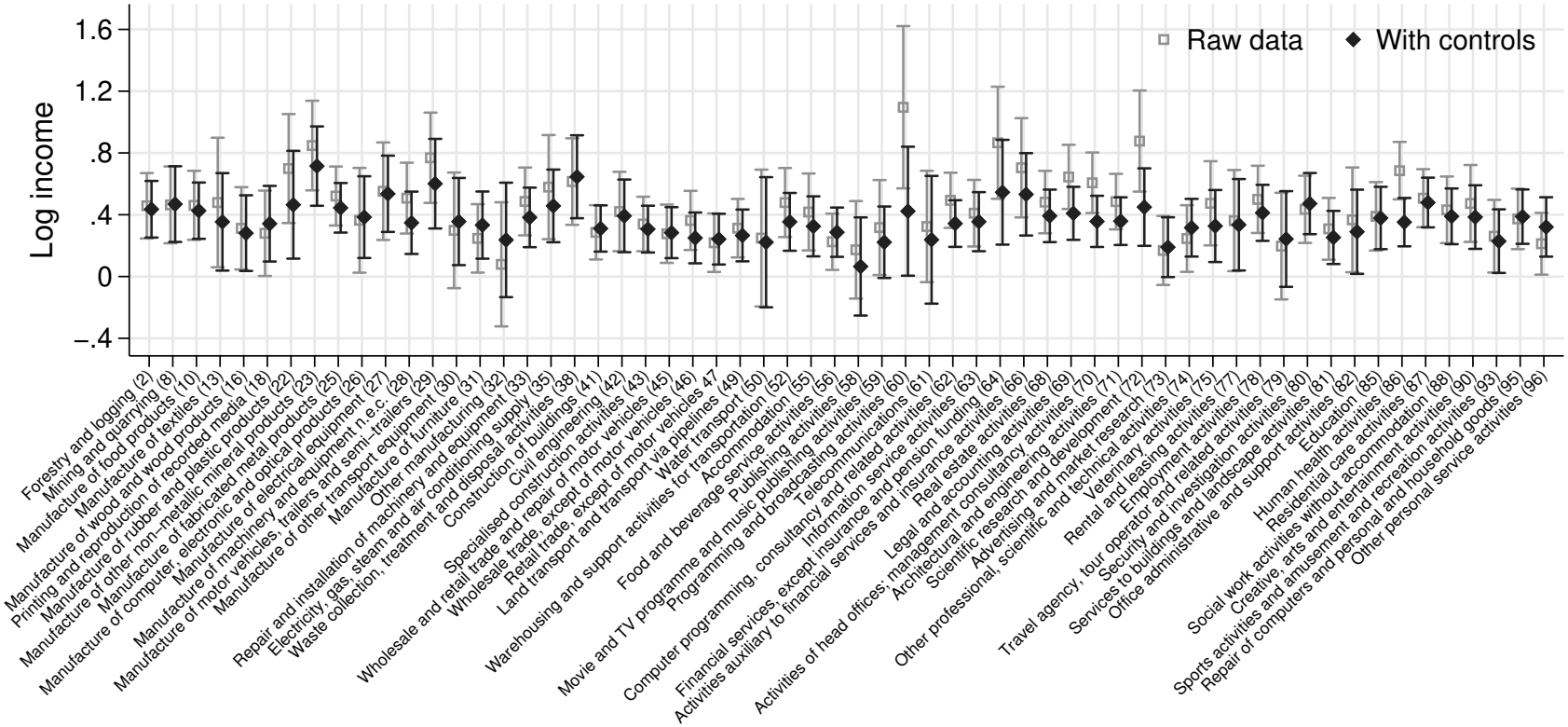
Notes: Table presents the average increase in disposable income among new incorporated entrepreneurs from one year before to 10 years after they established their first business using different subsamples of our data. First, we exclude those who had an unincorporated business before becoming incorporated. Second, we show the gains separately for those with and without entrepreneurial parents. Third, we separate entrepreneurs by whether they still are or are not the main owners of the business five years afters the business was established. Finally, we show how the gains differ by years and by how we identify incorporated entrepreneurs from the data. Importantly, the last two rows in the table highlight that when using the same time period (2006–2014), our findings on the gains from entrepreneurship do not differ depending on whether we define the ownership status based on Tax Administration data or the Statistics Finland firm ownership database.

Figure B2: Disposable income trajectories using a balanced panel of individuals



Notes: The figure presents the unconditional log disposable income trajectories of incorporated business owners and wage earners (pseudo-starters) relative to the year $t-1$ (denoted by zero in the figure) using three samples. Panel A shows our baseline unbalanced panel results. Panel B includes the panel data results for individuals we observed from $t-1$ to $t+10$. Panel C shows the panel data results for individuals observe from $t-5$ to $t+10$. Overall, the figure shows that the gains from entrepreneurship are larger when using a balanced panel data instead of an unbalanced panel as in our baseline analysis.

Figure B3: Gains in disposable income by industry of the newly established firm



Notes: The figure shows the individual-level gains in log disposable income by the industries of newly established incorporated firms. The gains are measured as the difference in annual income 10 years after and one year before the firms are established. The gains are presented relative to the industry with the lowest gains (manufacturing of chemicals and chemical products). Raw data corresponds to a model without controls. In the model with controls, we account for all interactions of age, sex and education (primary education, secondary degree or tertiary degree), these interactions using age squared and age cubed, and individual $(t-1)$ and parental income ranks. We use the TOL 2008 classification of Statistics Finland, which is based on the EU's classification of economic activities (NACE), to classify the industries.

C Taxation of incorporated and unincorporated firms in Finland

Sole proprietors and partnerships are pass-through entities, meaning that their profits are taxed only at the owner level as personal income. In contrast, privately held corporations are separately tax-liable, meaning that their profits are taxed at the firm level at the corporate tax rate. Owners of privately held corporations pay an additional tax on income withdrawn from the firm.

Finland applies a dual income tax system where earned income (wages, pension income etc.) and capital income (interest income, rental income, dividends, capital gains etc.) are taxed with separate tax schedules. The earned income tax rate schedule is more progressive, with a higher top tax rate of approximately 55%, whereas the tax rate for capital income is 30% for income below 30,000 euros and 34% for income above that threshold (in 2023).

Within the dual income tax system, the declared profits of unincorporated firms are divided into earned income and capital income components based on the net assets of the firm (assets minus liabilities), such that the amount corresponding to 20% of the net assets is taxed as capital income and any remaining profit as owners' earned income.[17]

The tax schedule for incorporated firms is more complex. Privately held corporations pay a 20% tax on profits, and the owners pay earned income tax on any wage income withdrawn from the firm (wages are deductible from firm profits). In addition, corporations can distribute dividends to their owners, which are in general taxed as the owners' personal capital income. However, the dividend income tax schedule for privately held corporations is more complicated, including one further tax rate kink determined by firm-level net assets (8% of net assets) and another kink based on the euro amount of the dividends withdrawn from the firm (150,000 euros). The tax rate increases for dividend income above these thresholds, so as to suppress the incentives to minimize personal income taxes by shifting income from the earned income tax base to the more leniently taxed capital income tax base.

In more detail, the tax rate for dividend income that falls below the amount corresponding to 8% of a firm's net assets is taxed at a flat effective tax rate of 26%, including both owner-level dividend taxes and corporate taxes. This dividend income is 75% tax-free, and 25% is taxed as personal capital income. Combined with the corporate tax of 20%, this yields an effective tax rate of 26% (0.20+(0.80*0.25*0.30)=0.26). The rate increases to 26.8% if the annual personal capital income of the owner exceeds 30,000 euros. The tax rate

[17]The owners can also choose a smaller capital income component of either 10% or 0%. These choices can be preferable for unincorporated firms with small profits in which case the earned income tax rate of the owner(s) can be below the personal capital income tax rate (30% or 34%).

increases to 40.4% (43.1% if capital income above 30,000 euros) for dividends below the net assets threshold that exceed 150,000 euros. Dividend income that exceeds the 8% net assets threshold is partly tax-free (15%) and partly taxed as earned income (85%) according to the progressive wage tax schedule excluding social security contributions. The exact tax rates and thresholds have varied over time in the time period we study, but the baseline system has remained the same.[18]

[18]See Harju et al. (2022) for details on the recent reforms and Harju and Matikka (2016) for more details on the dividend taxation of privately held corporations.

D Definitions

In this section, we present the definitions of the samples and variables we use in our baseline analysis.

Estimation Sample: We restrict our baseline sample to individuals who are 25–64 years old when establishing their business and to individuals with positive annual wage income and/or entrepreneurial income. We also restrict our main sample to those individuals for whom we observe child-parent links. This last restriction eliminates all individuals born before 1953. We analyze firm starts and pseudo-starts (see wage earners below) that take place in 1998–2014.

Incorporated entrepreneurs, IE: Main owners of newly established privately held corporations. An individual is classified as a new incorporated entrepreneur in the year when his/her firm shows positive sales if the individual had no prior ownership in other privately held corporations. Data on full ownership information from Statistics Finland starts from 2006 (Finnish Longitudinal Owner-Employer-Employee Data FLOWN). For the period 1997–2005, we rely on data from the Finnish Tax Administration that cover owners who received dividend income from their firm.

Unincorporated entrepreneurs, UIE: Sole proprietors or main owners of newly established partnerships. An individual is classified as an new unincorporated entrepreneur in the year when the firm shows positive sales and the owner obtains more than 50% of his/her market income as business income, if the individual had not been a sole proprietor before and had no prior ownership in other partnerships. The ownership data start from the year 1997 and are from the Finnish Tax Administration.

Wage earners, WE: We label as wage earners individuals who are 25–64 years old and have positive annual wage income and who have no ownership of any firm (incorporated or unincorporated) over the whole time period we observe in the data. To these individuals we assign a pseudo-start year ($t = 0$) that is drawn randomly from a uniform distribution. These wage earners do not actually start a business that year, but this approach enables us to benchmark the income development of new business owners against the income growth of wage earners over time.

Disposable income: Sum of earned income, entrepreneurial income, capital income and transfers and benefits received after taxes and other levies.

Market income: Sum of earned income, entrepreneurial income and capital income before taxes and other levies.

Retained earnings: We follow a similar approach as in Alstadsæter et al. (2016) and allocate earnings retained in the firm instead of distributed as dividends to the owners to each owner of an incorporated firm based on their ownership share of the firm. We include this income on top of the market income observed in the administrative data to study the significance of retained earnings in income development over time.

Income rank: Individual's position in the income distribution in a given year for a given income concept. Depending on the scope of the analysis, we use either percentile (1–100), decile (1–10) or quintile (1–5) ranks.

Parental income: Average annual household market income of parents when they were 45–50 years old. Parental income rank refers to household's position in the parental household distribution. We cover parents who are 45–50 years old after 1987. As a sensitivity check, for a subset of parents who are 45–50 years old after 1995, we can define parental ranks using disposable income and capital income.

Parental entrepreneurship status: Parents are classified as entrepreneurs if at least one of the parents is labeled as a business owner for at least 5 years according to the socioeconomic status defined by Statistics Finland. This approach enables us to use a similar definition for parental entrepreneurship using our full data starting from 1987.

Education: The educational level and field of the highest qualification/degree. Education classifications correspond to the ISCED 2011 classification.

Firm entry: First year of a business is defined as the first year the firm is both registered at the Finnish Patent and Registration Office and presents positive sales in the tax return data.

Firm ownership: For the owners of privately held corporations, we use two separate data sets to define firm ownership: 1) the main owner information from the Tax Administration for those owners who received dividend income from their firm, available for 1997–2016, and 2) the full ownership database from Statistics Finland, which is based on information from the Finnish Patent and Registration Office, available for 2006–2019. We use a data set including all sole proprietors and partnership owners from the Tax Administration available

for 1997–2016.

Main Ownership Status: We define the main owner of an incorporated business as the owner who owns the largest share of the firm.

Bankruptcy: A firm is defined as bankrupt if bankruptcy is declared by the district court, following the bankruptcy regulations in Finland.

Initial equity: Invested equity of the firm in the first year after it was registered at the Finnish Patent and Registration Office.

Sales: Annual sales income from products and services after taxes.

Number of employees: The number of full-time equivalent employees in the firm during a year.

Staff costs: The annual sum of wages, salaries and personnel expenses determined on the basis of wages or salaries, such as pension contributions and social security contributions, but excluding income taxes.

Value added: Annual value of sales minus variable costs. Variable costs include intermediate inputs used in production, such as materials and services.

Productivity: Value added divided by the number of employees.

R&D investments: Annual investments in research and development. In our analysis, we use an indicator variable of whether or not a firm has positive R&D investments.

E Overview of data used in previous related studies

Previous studies on the link between entrepreneurship and personal income have relied on surveys using either cross-sectional data or panels with relatively few individuals followed over time. Hamilton (2000), who documents that self-employment is associated with lower initial earnings and lower earnings growth compared to paid employment, relies on a sample from the Survey of Income and Program Participation that covers 8,771 males observed in 1983–1986 in the US. Both incorporated and unincorporated business owners are labeled as self-employed if they report self-employment as the main "non-casual" labor market activity for at least three months in a given 12-month spell. Moskowitz and Vissing-Jørgensen (2002) show that, in the US, ownership in privately held businesses is highly concentrated, yet the returns to this poorly diversified investment are at par with stock market returns. Their main source of data is the 1989, 1992, 1995 and 1998 Survey of Consumer Finances, which covers roughly 4,000 households per survey year.

Levine and Rubinstein (2017) make a seminal contribution by disaggregating the self-employed into incorporated and unincorporated. Their first data source is the Current Population Survey, which they use as cross-sections and as a two-year panel. In the pooled cross-sections, there are approximately 40,000 incorporated and 75,000 unincorporated business owners. Their second source is the National Longitudinal Survey NLSY79, which enables the authors to follow roughly 2,000 incorporated and 9,000 unincorporated business owners over 1982–2012.

We depart from these previous studies by having highly detailed administrative population-wide panel data on individuals and firms, including the exact establishment dates and ownership structures of new businesses. These data allow us to follow the same individuals over time before and after they become business owners for the first time in this literature. Moreover, we have access to detailed information about entrepreneurs, their parents and their newly established businesses. The study closest to ours is that by Berglann et al. (2011), who follow a sample of 12,000 new entrepreneurs, both incorporated and unincorporated, who start their business between October 2000 and October 2001 in Norway five years before and five years after they become entrepreneurs.

Halvarsson et al. (2018) and Lindquist and Vladasel (2022) have similar Swedish administrative data but a different scope. Halvarsson et al. (2018) examine how self-employment shapes the income distribution using two cross-sections of data (2005 and 2013). Lindquist and Vladasel (2022) construct lifetime earnings for entrepreneurial sons and their fathers to document that, driven by selection into entrepreneurship, incorporated entrepreneurs are more upwardly mobile than wage earners.