DIW BERLIN

1683

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2017

Media and Occupational Choice

Alexander Konon and Alexander Kritikos

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Media and occupational choice*

Alexander Konon[†], Alexander Kritikos[‡]

August 2017

We address the question of whether media influences occupational choices. To theoretically examine media effects, we construct a dynamic Bayesian occupational choice model with sequential decisions under ambiguity due to imperfect information. We show that sufficiently intensive positive media articles and reports about entrepreneurship increase the probability of self-employment and decrease the probability of wage work. To test our model, we use an instrumental variable approach to identify causal media effects using US micro data and a country-level macro panel with two different media variables. We find that an increase in positive media articles and reports about entrepreneurs generates effects on choice probabilities that are consistent with our model.

JEL classification: D81, D83, J62, L26
Keywords: Media, occupational choice, Bayesian learning, ambiguity aversion

*We are indebted to Sascha O. Becker, Yves Breitmoser, Arnaud Chevalier, and Hani Mansour for their helpful comments.
[†]Alexander Konon is Research Associate at the German Institute for Economic Research (DIW Berlin), e-mail: akonon@diw.de.
[‡]Alexander Kritikos is Research Director at the German Institute for Economic Research (DIW Berlin), Professor of Economics at the University of Potsdam, and Research Fellow of the IZA, Bonn, and of the IAB, Nuremberg, e-mail: akritikos@diw.de.

1 Introduction

Research shows that media, often recognized as the fourth branch of government in democracies, has a significant impact on a number of economic and political variables. For instance, by blaming persons or institutions for violating certain rules or shaming them for being underperformers, the media is able to play a corporate governance role for publicly traded companies (Zingales 2000; Dyck, Volchkova & Zingales 2008). Media also influences voting behavior (Della Vigna & Kaplan 2007) and voter turnout (Gentzkow 2006). In general, media can shape the image of public figures, institutions, groups, and individuals.

One group of individuals attracting a significant amount of media attention in recent times are entrepreneurs; including reports in many countries that celebrate entrepreneurs as "heroes."[1] In this contribution, we ask whether media articles and reports about entrepreneurs have an effect on the occupational choices of media consumers. Thus, we are not interested in the direct effect on those who were placed on the pedestal but in the indirect effects channeled through media consumption. To be more specific, we analyze, from a theoretical and empirical point of view, whether media articles and reports about successful entrepreneurs, which are unlikely to change actual probabilities to succeed in self-employment (the distribution of outcomes) but may change beliefs (the distribution of subjective outcome probabilities), influence occupational choice decisions.

To theoretically examine potential effects of positive media articles and reports about entrepreneurs, we construct a dynamic occupational choice model with Bayesian learning. In our model, individuals select an occupation given that they are only imperfectly informed—the outcomes of their choices are subject to ambiguity. We compare optimal choices to choices of ambiguity-averse individuals, where ambiguity aversion may vary across occupational options. We show that, if ambiguity aversion associated with self-employment is higher than aversion linked to wage work, there is a bias against self-employment, in the sense that self-employment is selected with a lower than optimal probability.

By assuming that individuals making occupational choices also use information from the media, which is accessible at negligibly low cost, we integrate positive media reports about entrepreneurship as informational shocks. Based on our model, we derive two predictions. We establish that sufficiently intensive positive media reports about entrepreneurs increase the likelihood to select self-employment, while the probability to select wage work is reduced.

The two predictions are empirically tested with two different data sets. The first data set (we refer to it as the "micro panel") is based on the US National Health Interview Survey, providing rich individual-level information on occupational status and income, as well as on various demographic and socio-economic characteristics. Media consumption of positive articles and reports about entrepreneurs is approximated by the regional frequency of the search item 'famous entrepreneurs' provided by the Google Trends tool. We use the number of natural disasters, from the International Disaster Database, in non-US regions to introduce an exogenous variation in the media variable. Natural disasters in non-US regions represent natural-experiment-type exogenous shocks that are not related to driving factors of occupational choice in the US but

[1] For example, in March 2009, the *Economist* presented a special issue on entrepreneurs under the title "Global heroes." Along the same lines, the German newspaper *Frankfurter Allgemeine Sonntagszeitung* published specials on the start-up scene in Berlin (in June 2014 and August 2015) and Tel Aviv (in October 2015). There are also entrepreneurship-related television series such as *How I Made My Millions* (CNBC channel), *CNBC Titans* (CNBC channel), as well as *Shark Tank* (on the ABC network in the USA) and its German adaptation *Die Höhle der Löwen* (VOX channel), which portray a variety of successful entrepreneurs in different situations. There are many other positive reports about entrepreneurs all around the world.

affect media reports, as disasters generate top-priority news. Taking into account potential heteroskedasticity, which, if unaccounted for, leads to inconsistent estimates, effect directions are identified with a heteroskedastic IV probit approach.

The second data set (the "macro panel") is a self-constructed country-level panel using, *inter alia*, data on choice frequencies from the World Bank and data on media reports from the Global Entrepreneurship Monitor (GEM). In the macro panel, media consumption of positive articles and reports about entrepreneurs is approximated by the working-age-population share of individuals noticing frequent reports about successful entrepreneurs, provided by the GEM. Natural disasters in other countries are used to construct a sufficiently strong instrument for the media variable in the second data set. We estimate instrumental variable regressions for the probability of self-employment and wage work. In both data sets, including two different media variables, we are, thus, capable to identify causal effects.

Based on two IV regressions, we find, in support of our hypotheses, that positive media reports about entrepreneurs significantly increase the probability of selecting self-employment and reduce the probability of wage work. Using linear probability models, our macro panel also allows us to approximate effect sizes. We find that a one percentage point increase in the media variable increases the probability of self-employment by 0.5 percentage points and decreases the probability of wage work by 0.4 percentage points. Effect sizes are consistent with previous findings on persuasion effects in the literature.

Overall, we contribute to the existing literature in two ways. First, we develop an occupational choice model operating under ambiguity that allows for a direct assessment of media effects on choice probabilities. Secondly, we provide first empirical evidence for media effects on occupational choices.

The remainder of the paper is organized as follows. In Section 2, we review previous research related to our approach. Section 3 presents the theoretical model. In Section 4, we analyze the effects of media on occupational choices in the theoretical model. Section 5 provides empirical results. In Section 6, we summarize and conclude. The Appendix contains proofs, additional information, and supplementary results.

2 Previous research

This section provides an overview of two strands of research, which we aim to combine and extend. First, we provide a brief review of the existing empirical evidence on the impact of media on individual and institutional behavior. Second, as we build upon research on decisions under ambiguity, we discuss theoretical concepts related to the so-called multi-armed bandit problem, which is a simple way to model decisions with ambiguous outcomes. We, then, outline our research approach.

2.1 Research on media impact

Recent research shows that media significantly impacts the decisions of individuals and institutions. For instance, Dyck et al. (2008) reveal that media shaming of corporate governance violations increases the probability of their reversal in Russia. Della Vigna & Kaplan (2007) establish that the conservative Fox News Channel convinced a substantial share of its non-Republican viewers to vote Republican in US presidential elections between 1996 and 2000. Enikolopov, Petrova & Zhuravskaya (2011) find that media impacts voting behavior in Russia. Gentzkow (2006) establishes a negative effect of television on voter turnout in the United States. Dyck, Moss & Zingales (2008) demonstrate that media coverage of certain topics (such as poverty) can affect the voting behavior of US Congressional Representatives and Senators (also, see Besley & Burgess 2002; Strömberg 2004; Eisensee & Strömberg 2007).

2.2 Modeling decisions under ambiguity with multi-armed bandits

Individuals making occupational choices are assumed to operate in an *ambiguous* environment, where we use the standard definition of ambiguity caused by imperfect information (Fellner 1961; Frisch & Baron 1988; Camerer & Weber 1992):

> Ambiguity is uncertainty about probability, created by missing information that is relevant and could be known. (Camerer & Weber 1992, p. 330)

As an alternative, the literature also uses the notion of 'uncertainty' (Camerer & Weber 1992). According to the definition above, uncertainty is synonymous with an ambiguous choice environment.[2]

Occupational choices are characterized by ambiguity of choice outcomes. For instance, the launch of a new product by an entrepreneur is usually associated with ambiguity regarding market reaction. Consequently, theoretical and empirical models in the literature, such as Jovanovic (1979), MacDonald (1988), Hintermaier & Steinberger (2002), Vereshchagina & Hopenhayn (2009), Campanale (2010), Poschke (2013), and Manso (2016), treat starting a business as an experiment with an unknown outcome. The dynamics of wage growth are also consistent with the assumption that information about unknown workers' skills is only gradually revealed to the employer (Antonovics & Golan 2012). Thus, career decisions are not final but rather a process of trial and error with learning resulting in transitions between occupations.

By allowing to model different types of sequential decisions under ambiguity, the bandit problem, a fairly general framework, accounts for the most prominent features of occupational choices, beyond a deterministic or risky choice environment. The conventional description of the general $\mathfrak{D}$-armed bandit problem is as follows. Assume that, in a casino, there are $\mathfrak{D} \in \mathbb{N}$ one-armed bandits that can be played by a gambler. Pulling one arm results in a reward generated by some distribution that is usually unknown, making the decision environment ambiguous. However, pulling one arm and observing the outcome provides information about the underlying reward distribution, such that the gambler can learn. Reward distributions are usually assumed to be different across the $\mathfrak{D}$ arms but there may exist dependencies between them. Given some time horizon and an objective function (for instance, the expected sum of rewards), the gambler must decide which of the $\mathfrak{D}$ arms to play; how many times to play each arm; and in which order to play them. Gittins, Glazebrook & Weber (2011) provide an extensive overview on Bayesian multi-armed bandit problems and corresponding problem solutions.

Rothschild (1974) uses a two-armed bandit to analyze equilibrium price distributions given that firms have imperfect knowledge about demand functions and need to experiment to find the profit-maximizing price. Jovanovic (1979) examines employee turnover as a consequence of learning processes. Bergemann & Hege (2005) use bandits to examine best financing rules for research projects with unknown length and success. Antonovics & Golan (2012) investigate career patterns of workers under the condition that skills are unknown. Konon (2016) examines how ambiguity preferences influence occupational choices under ambiguity.

[2] Note that this statement does not hold for Knightian uncertainty (Knight 1921), which is immeasurable in principle.

2.3 Our research approach

In this contribution, we combine the two strands of research above. To the best of our knowledge, we are the first to theoretically and empirically analyze the impact of media on occupational choices.

In our model, we use a two-armed bandit with ambiguous arm-specific reward distributions, Bayesian learning, and a joint prior distribution for reward probabilities. Our formulation of the problem builds upon the work of Bradt, Johnson & Karlin (1956) and Konon (2016). In such a setting with ambiguity, even rather unspecific information from the media is theoretically able to affect choices by influencing beliefs about how likely success in self-employment is in general. Put differently, media—for instance, an article about a famous entrepreneur—might not be able to affect the distribution of outcomes, to change risk, but it may change the meta distribution of outcome probabilities constituting the beliefs of an individual.[3]

The distinctive feature of our model is that, given our assumptions on the prior distribution and outcome distributions, we are able to derive simple expressions for the distributions of choices. Consequently, we can directly examine properties of theoretical choice probabilities, whereas most of the aforementioned papers using bandits only derive general properties of strategies. This allows us to compare theoretical choice distributions to observable choice probabilities.

Based on a proposition resulting from our theoretical model, we derive two hypotheses allowing us to examine the central question of this paper: the question of whether media influences occupational choices. The hypotheses are tested with two different data sets, a micro panel from the US that allows investigating individual behavior and a macro panel combining information from a larger number of countries. Using a micro and a macro panel allows us to provide results on media effects within a country and between-country effects. This is in contrast to most studies evaluating media effects that either concentrate on cross-country differences or differences (in media coverage of certain topics) within a country (Dyck et al. 2008). Moreover, our macro panel includes both developed and a relatively large number of developing economies; the latter being underrepresented in the analysis of media effects.

3 A model of occupational choices under ambiguity

To theoretically analyze under what conditions articles and reports affect occupational choices, we construct an occupational choice model with ambiguity and learning. Choices are driven by decision rules. In the following, we differentiate between two potential decision rules. The first rule maximizes expected success outcomes and, thus, allows deriving individually optimal behavior. The second decision rule is based on the assumption that individuals dislike ambiguity (Ellsberg 1961) and, as they prefer to avoid ambiguous situations, they may decide for options that are not optimal in terms of income expectations but better in terms of ambiguity avoidance. The section first introduces the model's setup, followed by a discussion of the basic assumptions of the model, and an examination of decision rules and individual behavior.

[3]Consuming media articles and reports about successful or famous entrepreneurs cannot affect choices in a setting with deterministic outcomes and perfect information by construction. Risk taking is an important factor underlying entrepreneurial activities (e.g., Djankov, Qian, Roland & Zhuravskaya 2006; Vereshchagina & Hopenhayn 2009), but the reading of an article about a famous and successful entrepreneur is unlikely to reduce the entrepreneurial risk of an individual, as information provided by the article is not specific enough to be relevant for the individual's future business. Consequently, even in a setting with stochastic outcomes but known outcome distributions—a setting with risk—media is unlikely to have any effect.

3.1 Setup

There are two occupational options an individual can choose from: self-employment $\mathcal{S}$ and paid work $\mathcal{W}$. Let $\mathbb{O} \equiv \{\mathcal{S}, \mathcal{W}\}$ denote the set of all available options. In reality, there is the alternative of unemployment as well. However, an active choice of voluntary unemployment cannot be identified in the micro and macro data we use, and is hard to identify in data in general. Therefore, we restrict our attention to self-employment and wage work. Yet, the model can be easily extended to account for unemployment.

Every option in $\mathbb{O}$ is associated with an i.i.d. reward sequence $\{\Omega_{i,n}\}_{n=1}^{N}$, where after a fixed and known period $N > 1$ the individual retires. Each reward sequence is based on a reward distribution F, such that $\Omega_{i,n}$ is generated by F with an option-specific parameter ϕ_i. Rewards come in form of occupational successes and failures, where $\omega = 1$ represents a success and $\omega = 0$ a failure (henceforth Assumption 1, discussed further below). Thus, reward distributions are Bernoulli (F is Bernoulli) and ϕ_i is the probability of succeeding in occupation $i \in \mathbb{O}$. A success is generated with probability ϕ_i and a failure occurs with probability $1 - \phi_i$.

We impose the following restrictions. For the probabilities to succeed in wage work and self-employment, we have $\phi_{\mathcal{W}}, \phi_{\mathcal{S}} \in (0, 1)$. Furthermore, we assume that

$$\phi_{\mathcal{W}} + \phi_{\mathcal{S}} = 1 \tag{1}$$

such that individuals may either be successfully self-employed or successful wage workers, but success probabilities in wage work and self-employment are almost never the same. This assumption (henceforth Assumption 2) is based on Lazear (2005) and is discussed in the next subsection. Individuals know that their probability to succeed in self-employment is decreasing in the probability to succeed in wage work, i.e., (1) is common knowledge.

Information is assumed to be imperfect. Thus, the probabilities to succeed in wage work and self-employment are both unknown, implying ambiguity (henceforth Assumption 3, justified below). However, individuals have some prior knowledge. Furthermore, individuals obtain additional information about an option $i \in \mathbb{O}$ by selecting it and observing the outcome, reward drawn from $F(\phi_i)$.

Prior knowledge is given by successes in wage work $a_{\mathcal{W},0} \in \mathbb{N}^+$ and self-employment $a_{\mathcal{S},0} \in \mathbb{N}^+$ that individuals draw from a set of historical data. Historical data can be represented by reward observations of other individuals, such as parents and peers[4] (e.g., Minniti 2005; Bosma, Hessels, Schutjens, Van Praag & Verheul 2012[5]).

Prior distributions are Dirichlet. The Dirichlet distribution is a proper conjugate prior for probabilities, where the condition $\phi_{\mathcal{S}} = 1 - \phi_{\mathcal{W}}$ holds, and has density

$$\varphi(\phi_{\mathcal{S}}, \phi_{\mathcal{W}}; a_{\mathcal{S},n}, a_{\mathcal{W},n}) = \check{\Gamma}(a_{\mathcal{S},n}, a_{\mathcal{W},n}) \phi_{\mathcal{S}}^{a_{\mathcal{S},n}-1} \phi_{\mathcal{W}}^{a_{\mathcal{W},n}-1}, \quad \check{\Gamma}(x_1, x_2) \equiv \frac{\Gamma(x_1 + x_2)}{\Gamma(x_1)\Gamma(x_2)} \tag{2}$$

$a_{\mathcal{S},n}$ and $a_{\mathcal{W},n}$ are parameters of the distribution, and Γ is the gamma function. Given no actual observations of rewards but some set of historical data, a success probability ϕ_i for $i \in \mathbb{O}$ obeys

[4]Information from parents, spouses, and peers can either encourage or discourage a certain occupational choice. For instance, there is anecdotal evidence that wives and parents tend to block the pursuit of entrepreneurship in Japan, while American parents tend to encourage entrepreneurial activities (Fifield 2016).

[5]Note, however, that Bosma et al. (2012) use the concept of role models, which is much richer than our concept of information because besides information role models also provide support and guidance.

the following distribution:

$$\varphi(\phi_i; a_{\mathcal{S},0}, a_{\mathcal{W},0}) = \check{\Gamma}(a_{\mathcal{S},0}, a_{\mathcal{W},0})\phi_i^{a_{i,0}-1}(1-\phi_i)^{a_0-a_{i,0}-1}, \quad a_0 \equiv a_{\mathcal{S},0} + a_{\mathcal{W},0} \tag{3}$$

Actual observations (i.e., non-historical data) change prior distributions according to Bayes' law. Assume that in some period $n > 0$ wage work is selected and the reward $\omega_{\mathcal{W},n} \in \{0, 1\}$ is observed. Successes in wage work until period n are given by $a_{\mathcal{W},n-1}$. Successes in self-employment are given by $a_{\mathcal{S},n-1}$. Then, the posterior distribution (given a Dirichlet prior, the posterior is also Dirichlet) of the probability to succeed in wage work is

$$\varphi(\phi_{\mathcal{W}}; a_{\mathcal{S},n}, a_{\mathcal{W},n}) = \check{\Gamma}(a_{\mathcal{S},n}, a_{\mathcal{W},n})\phi_{\mathcal{W}}^{a_{\mathcal{W},n}-1}(1-\phi_{\mathcal{W}})^{a_n-a_{\mathcal{W},n}-1} \tag{4}$$

where $a_{\mathcal{W},n} = a_{\mathcal{W},n-1} + \omega_{\mathcal{W},n}$, $a_{\mathcal{S},n} = a_{\mathcal{S},n-1} + 1 - \omega_{\mathcal{W},n}$, and $a_n \equiv a_{\mathcal{S},n} + a_{\mathcal{W},n}$. The posterior distribution of the probability to succeed in self-employment is obtained in a similar way.

The general setup of the model, established above, introduces a sequential decision problem: Individuals have to decide which occupation to select in every period $n = 1, \ldots, N$. Let $\mathfrak{d}_n \in \mathbb{O}$ denote the decision in period n. We assume that individuals use a decision rule that determines the probabilities to select an option. In other words, a decision rule generates $\mathbb{P}(\mathfrak{d}_n = i)$ for all $i \in \mathbb{O}$ and all n.

The analysis of the model consists of two steps. First, we demonstrate that sufficiently high ambiguity aversion generates choice probabilities that differ from the optimal probability maximizing the expected sum of individual successes. Second, we show that media can change choice probabilities and, in particular, media is able to reduce or even eliminate the difference between optimal and ambiguity-aversion-affected probabilities. However, before analyzing decisions, we first discuss modeling assumptions.

3.2 Assumptions

Our model's setup rests on three basic modeling assumptions. The first assumption determines the type of rewards by restricting it to successes and failures. The assumption is helpful for two reasons. First, it simplifies modeling. Second, as it is relatively easy to find or construct a measure of the number of media reports on successful entrepreneurs, it makes it possible to conduct an empirical analysis.

Assumption 1. Occupational options produce rewards in form of periodical occupational successes or failures according to some distribution. Occupational options may differ with respect to their ability to deliver successes such that reward distributions can be different across options.

Occupational successes can be defined in various ways. A simple definition is that a success is achieved when an individual reaches a self-set monetary income benchmark. More formally, let Π denote the monetary reward generated by an arbitrary occupation. Let F_Π denote the corresponding continuous distribution function of monetary rewards. Furthermore, let B_Π denote a self-set income benchmark. A success occurs if the monetary income is above the benchmark. Consequently, the probability of a success is

$$\phi = \mathbb{P}(\Pi > B_\Pi) = 1 - \int_{-\infty}^{B_\Pi} f_\Pi(\pi)\mathrm{d}\pi$$

Our model is constructed under the assumption that success probabilities are unknown. This is fully consistent with the definition above if we assume that the distribution of monetary rewards, F_Π, is unknown, such that $\mathbb{P}(\Pi > B_\Pi)$ cannot be directly computed.

The type of rewards fixed by Assumption 1 does not necessarily contradict the standard way to assess rewards or incomes using the expected value—an option is "better" if it yields a higher expected income. The following example demonstrates this conjunction for the most common distribution of incomes: the log-normal (see Lopez & Servén 2006).

Example 1. Assume that Π has a log-normal distribution such that $\mathbb{E}[\log \Pi] = \mu_\Pi$ and $\mathbb{V}[\log \Pi] = \sigma_\Pi^2$. Assume that there are two options where $\mu_{\Pi,1} > \mu_{\Pi,2}$, while $\sigma_{\Pi,1} = \sigma_{\Pi,2} = \sigma_\Pi$. As $\mathbb{E}[\Pi] = \exp(\mu_\Pi + \sigma_\Pi^2/2)$, option 1 generates a higher expected income than 2. Let the benchmark be sufficiently large such that $B_\Pi > 1$ (for instance, larger than one unit of money). The success probability of an arbitrary option is $\phi = 1 - \mathbb{P}(\Pi \leqslant B_\Pi) = 1/2 - 1/2\,\mathrm{erf}([\log B_\Pi - \mu_\Pi]2^{-\frac{1}{2}}\sigma_\Pi^{-1})$ where erf is the Gauss error function. Hence, we get

$$\phi_1 - \phi_2 = \frac{1}{2}\left[\mathrm{erf}\left(\frac{\log B_\Pi - \mu_{\Pi,2}}{\sqrt{2}\sigma_\Pi}\right) - \mathrm{erf}\left(\frac{\log B_\Pi - \mu_{\Pi,1}}{\sqrt{2}\sigma_\Pi}\right)\right]$$

Using the properties of the error function, it is easy to show that $\phi_1 - \phi_2 > 0$ for $B_\Pi > 1$. Consequently, $\mathbb{E}[\Pi_1] > \mathbb{E}[\Pi_2]$ transforms into $\phi_1 > \phi_2$.

The second assumption establishes how success probabilities are related.

Assumption 2. Individuals are either productive in self-employment or in paid employment but almost never both at exactly the same level.

Lazear (2005) theoretically and empirically shows that the self-employed are rather jacks-of-all-trades than specialists (also, see Wagner 2006; Stuetzer, Goethner & Cantner 2012). Assumption 2 builds on this finding.[6] In our model, the probability of succeeding in wage work is implicitly assumed to increase in specialization. A specialist with much-needed skills will experience high rewards in wage work but low rewards in self-employment since highly developing a particular skill is not possible without neglecting all other skills.

Figure 1 explains how jacks-of-all-trades and specialists are related to each other.[7] Consider point "O." If the individual decides to specialize on one skill, she will increase her probability to succeed in wage work but simultaneously decrease her probability to succeed in self-employment. An even development of all skills will decrease the probability to succeed in wage work but increase the probability to succeed in self-employment. However, it is not possible to increase both probabilities at the same time.

The third assumption introduces imperfect information.

Assumption 3. The probabilities to succeed in wage work and self-employment are unknown.

There are several reasons for why reward distributions are unknown. The reward from entrepreneurship depends on many factors that individuals cannot control or fully anticipate.

[6]Lazear's (2005) approach is empirically tested but it does not capture some types of individuals. For instance, individuals who are strongly restricted in their choice of occupation, for example, due to severe poverty or disabilities, must be excluded from the analysis. Furthermore, a certain level of *basic* education is necessary for wage work and self-employment, such that the assumption holds conditional on basic education levels.

[7]The most important implication of Assumption 2 is that, for most individuals, one option (self-employment or wage work) is clearly better than the alternative. The assumption also results from Lazear's (2009) skill-weights approach, where skills are general but jobs of different types weight the same skill in different ways. From the perspective of skill-weights, different weights for entrepreneurship and wage work can result in different success probabilities.

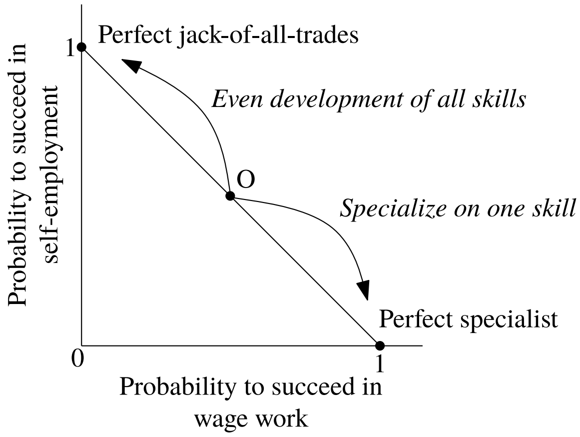


FIGURE 1. Jacks-of-all-trades and specialists

If an entrepreneur launches a new product, she may make a certain prediction about how the market will react to it, but there is still substantial ambiguity about market success—highly innovative products often tend to be rejected by the market.

The reward from wage work is unknown because wage workers do not have full control over their careers. The probability of losing a job or being promoted is usually not perfectly known. Furthermore, there is evidence that skills are also unknown. Antonovics & Golan (2012) show that patterns of occupational choices and wage growth are consistent with the assumption that jobs only gradually reveal information about unknown workers' skills.[8] Consequently, if some important skill influencing outcomes in wage work can only be revealed by actually doing some tasks, there will be ambiguity about outcomes and rewards.

3.3 Decisions: Rules and strategies

Given the setup depicted above, individuals are assumed to follow an occupational strategy based on a decision rule. As in standard economic theory, we assume that in every period individuals assign a measure of utility to every option in $\mathbb{O}$ and select the option with the highest utility. In the context of multi-armed bandits, researchers label such an approach as index strategy. An equivalent formulation is that individuals will decide based on relative utility. For instance, to decide between wage work and self-employment, individuals assign utility (index) $\mathsf{u}_{\mathcal{W}}$ to wage work and $\mathsf{u}_{\mathcal{S}}$ to self-employment, and decide for wage work if relative utility $\mathsf{u}_{\mathcal{W}} - \mathsf{u}_{\mathcal{S}}$ is weakly positive and for self-employment else.

We discuss two ways to formalize utility. The first approach is relative unbiased utility, which we use as a benchmark. The second approach is relative biased utility, which incorporates ambiguity preferences.

[8]Rewards of wage workers depend on their skills and the ability of employers to correctly assess these skills and set a corresponding wage. There is evidence that the productivity of young workers is an unknown variable for employers such that employers need some time to learn about the skills of their workers (e.g., Mansour 2012). Hence, ambiguity in wage work might be two-sided: In addition to ambiguity on the workers' side, there is a strong indication for ambiguity on the employers' side.

3.3.1 Unbiased utility

The idea behind the construct of relative unbiased utility is to select the most promising option, the option with the largest expected success probability, in every period by relying on priors and actual observations. A strategy exclusively concentrating on expected successes effectively ignores ambiguity because deviations from expected success probabilities are not assigned any relevance. Utility is unbiased because the only motive behind decisions is to always select the best option in expectations.

Relative unbiased utility is equivalent to a simple maximization of per-period expected rewards. In every period n, distributions of probabilities to succeed in wage work and self-employment are given by densities $\varphi(\phi_{\mathcal{W}}; a_{\mathcal{W},n}, a_{\mathcal{S},n})$ and $\varphi(\phi_{\mathcal{S}}; a_{\mathcal{W},n}, a_{\mathcal{S},n})$. Hence, given information in period n, expected probabilities to succeed are

$$\mu_{i,n} = \int_0^1 \phi_i \varphi(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}) \mathrm{d}\phi_i \quad \text{for } i \in \mathbb{O} \tag{5}$$

Equation (5) is updated in every period since either $a_{\mathcal{W},n-1} < a_{\mathcal{W},n}$ or $a_{\mathcal{S},n-1} < a_{\mathcal{S},n}$ (never both). The individual, then, selects the option that promises a success with the highest probability, while the option with the highest expected success probability may change as new information is obtained.

This decision rule corresponds to the following strategy: In period $n > 0$, use the expected success probability of wage work $\mu_{\mathcal{W},n-1}$ and the expected success probability of self-employment $\mu_{\mathcal{S},n-1}$ to construct $\mathfrak{u}_n = \mu_{\mathcal{S},n-1} - \mu_{\mathcal{W},n-1}$. Select self-employment if $\mathfrak{u}_n > 0$ and wage work if $\mathfrak{u}_n \leqslant 0$.

3.3.2 Biased utility

As demonstrated by Ellsberg (1961), individual decisions are not entirely based on expected outcomes but there is also a tendency to avoid ambiguity. For instance, assume that we have two options with exactly the same success probability, but the first option is more ambiguous than the second. It is reasonable to assume that an ambiguity-averse individual will exhibit a tendency to select the less ambiguous option over the more ambiguous one. Accordingly, decisions may not only be motivated by good performance in expectations but also by ambiguity avoidance. In such a case, relative utility can be biased because, besides good decision performance, preferences toward ambiguity also influence decisions.

Ambiguity can be defined as the variance of the distributions of success probabilities (Maccheroni, Marinacci & Ruffino 2013):

$$\mathsf{v}_{i,n} = \int_0^1 \{\phi_i - \mu_{i,n}\}^2 \varphi(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}) \mathrm{d}\phi_i \quad \text{for } i \in \mathbb{O} \tag{6}$$

An option $i \in \mathbb{O}$ is ambiguous in period n if $\mathsf{v}_{i,n} > 0$. Relative biased utility extends the idea of unbiased utility by simultaneously accounting for expected success probabilities and ambiguity.

Before introducing relative biased utility, it is necessary to discuss properties of preferences toward ambiguity. First, ambiguity aversion might be context-dependent. In our model, wage work and self-employment represent two different contexts. Wage workers are usually not directly responsible for covering damages. The self-employed do not have a buffer, in form of managers or employers, and must take the full responsibility for their actions (they have sufficient "skin in the game"). Put differently, self-employment can generate actual losses, while a wage is always non-negative, assuming that job loss corresponds to a zero wage. Hence, ambiguity in self-employment might be perceived as different (for instance, more problematic) than ambiguity in paid employment.

Second, Soto, John, Gosling & Potter (2011) show that individuals become more open to experience as they accumulate experience, a psychological effect consistent with the assumption that ambiguity aversion might be affected by (positive) experience. Third, and last, decisions tend to be self-reinforcing.[9] Positive psychological effects of successes constitute a simple approach to induce partially self-reinforcing decision patterns.

Hence, we postulate a last assumption establishing the existence of preferences toward ambiguity and their properties.[10]

Assumption 4. Individuals can be ambiguity-averse. Furthermore, ambiguity aversion can vary across occupations, and reactions to ambiguity may change over time depending on experience. However, the following conditions hold: (a) An individual who is ambiguity-averse never becomes ambiguity-affine or ambiguity-neutral. (b) In addition to informational effects, successes can have a psychological effect. A success can increase self-confidence. As a result, occupational-specific ambiguity aversion can effectively decrease.

The underlying idea of relative biased utility is as follows. Let $\mathrm{w}_{\phi-\mu}(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}, \theta_i)$ denote a function weighting potential deviations of actual success probabilities from the expected probability. We assume that the weight $\mathrm{w}_{\phi-\mu}$ depends on information $(a_{\mathcal{W},n}, a_{\mathcal{S},n})$ and a parameter $\theta_i \in \mathbb{R}$ representing option-specific preferences toward ambiguity. The utility of an option is given by

$$\eta_{i,n} = \mu_{i,n} + \xi_{i,n} \quad \text{for } i \in \mathbb{O} \tag{7}$$

$$\xi_{i,n} = \int_0^1 \{\phi_i - \mu_{i,n}\} \mathrm{w}_{\phi-\mu}(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}, \theta_i) \varphi(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}) \mathrm{d}\phi_i$$

The rationale behind Equation (7), which is similar to the mean-variance rule in portfolio choice (see Maccheroni et al. 2013), is that individuals will base their decisions on expected probabilities to succeed, represented by μ, but they will also anticipate potential mistakes, represented by $\phi - \mu$, which they might dislike, caused by the imperfect character of information.[11]

We use the following weighting function:

$$\mathrm{w}_{\phi-\mu}(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}, \theta_i) \equiv 1 - \frac{\theta_i}{a_{i,n}} \left\{ \phi_i - \mu_i(a_{\mathcal{W},n}, a_{\mathcal{S},n}) \right\} \quad \text{for } i \in \mathbb{O} \tag{8}$$

To verify that $\mathrm{w}_{\phi-\mu}$ is an appropriate weight, consider three types of preferences toward ambiguity.

[9] For the high explanatory performance of algorithms with choice-reinforcement components, see, for instance, Camerer & Ho (1999).

[10] A decision maker confronted with a risky option knows the distribution of the outcome (see, e.g., Holm, Opper & Nee 2013) and the known variance of the outcome can be used to measure risk (Tobin 1958). A decision in an ambiguous choice environment has to cope with the fact that outcome probabilities are unknown (Ellsberg 1961). This conceptual difference has an important implication with respect to the difference between risk and ambiguity preferences. Ambiguity preferences evaluate the distribution of outcome probabilities, which might change as new information is obtained, whereas risk preferences evaluate the known distribution of outcomes.

[11] Kahn & Sarin (1988) construct a similar representation—with a different weighting function—as an extension of subjective expected utility.

Ambiguity neutrality
Let $\theta_i = 0$ represent ambiguity neutrality. The weight is given by $\mathsf{w}_{\phi-\mu} = 1$. Hence:

$$\xi_{i,n} = \int_0^1 \phi_i \varphi(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}) \mathrm{d}\phi_i - \mu_{i,n} \int_0^1 \varphi(\phi_i; a_{\mathcal{W},n}, a_{\mathcal{S},n}) \mathrm{d}\phi_i = 0$$

Consequently, $\eta_{i,n} = \mu_{i,n}$ such that relative unbiased and biased utility are equivalent, i.e., decisions exclusively concentrate on expected success probabilities.

Ambiguity affinity
Let $\theta_i \in \mathbb{R}^-$ represent ambiguity affinity. Conditional on ambiguity affinity, success probabilities above the expected probability μ are assigned a weight larger than 1, while success probabilities below the expected success probability μ are assigned a weight smaller than 1. Thus, $\mathsf{w}_{\phi-\mu}$ emphasizes the following aspect of ambiguity: The true probability to succeed might be higher than the expected probability. More ambiguity will increase the utility of an option as

$$\frac{\partial}{\partial \mathsf{v}_{i,n}} \xi_{i,n} = -\frac{\theta_i}{a_{i,n}} > 0 \quad \text{for } i \in \mathbb{O} \text{ and all } n$$

Ambiguity aversion
Let $\theta_i \in \mathbb{R}^+$ represent ambiguity aversion. In this case $\mathsf{w}_{\phi-\mu}$ will emphasize negative estimation errors, i.e., the fact that the true probability to succeed might be smaller than the expected probability μ, by assigning a weight smaller than 1 if $\phi - \mu > 0$ and a weight larger than 1 if $\phi - \mu < 0$. More ambiguity also decreases the utility of an option as

$$\frac{\partial}{\partial \mathsf{v}_{i,n}} \xi_{i,n} = -\frac{\theta_i}{a_{i,n}} < 0 \quad \text{for } i \in \mathbb{O} \text{ and all } n \tag{9}$$

such that $\eta_{i,n} < \mu_{i,n}$.

Note that Assumption 4 holds since (9) never changes sign. An ambiguity-averse individual never becomes ambiguity-affine or ambiguity-neutral. Moreover, observed successes in an occupation have a self-confidence effect since $-\theta_i(a_{i,n}+1)^{-1} > -\theta_i a_{i,n}^{-1}$, such that given more successes individuals will react less negatively to more ambiguity. Finally, note that in consistency with Assumption 4, ambiguity preferences can be different across occupations since $\theta_i \in \mathbb{R}^+$ does not rule out $\theta_{\mathcal{S}} > \theta_{\mathcal{W}}$ (or $\theta_{\mathcal{W}} > \theta_{\mathcal{S}}$). For the remainder of the paper, we assume ambiguity aversion or $\theta_i \in \mathbb{R}^+$ for all $i \in \mathbb{O}$.

A strategy grounded in relative biased utility can be described as follows. In period $n > 0$, use the subjective utility of wage work $\eta_{\mathcal{W},n-1}$ and the subjective utility of self-employment $\eta_{\mathcal{S},n-1}$ to construct $\mathfrak{b}_n = \eta_{\mathcal{S},n-1} - \eta_{\mathcal{W},n-1}$. If $\mathfrak{b}_n > 0$, select self-employment. If $\mathfrak{b}_n \leqslant 0$, select wage work.

3.4 Individual behavior

A decision strategy induces a behavioral pattern. We assume that behavioral patterns are fully specified by the probabilities to select an option $i \in \mathbb{O}$ in some arbitrary period n. An important feature of our model, setting it apart from bandit models in the literature (e.g., Rothschild 1974; Jovanovic 1979; Bergemann & Hege 2005; Antonovics & Golan 2012; Konon 2016), is that it allows for the derivation of theoretical choice probabilities, which can, in principle, be directly compared to their empirical counterparts.

Behavior (probabilities to make a specific choice) induced by relative unbiased and biased utility is as follows.

Lemma 1. *Let $\mathfrak{d}^{\mathfrak{u}}$ denote a choice made by unbiased utility and $\mathfrak{d}^{\mathfrak{b}}$ a choice made by biased utility. Unbiased utility selects wage work with probability*

$$\mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{W}) = \mathbb{P}(\mathfrak{u}_n \leqslant 0) = H(\tau_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}}), \quad \tau_n^{\mathfrak{u}} = -\frac{n + a_{\mathcal{W},0} - a_{\mathcal{S},0}}{2}$$

and self-employment with probability $\mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{S}) = 1 - H(\tau_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}})$, where $H(x; n, \phi)$ is the cumulative distribution function of the binomial distribution given period n and success probability ϕ. Biased utility selects wage work with probability

$$\mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{W}) = \mathbb{P}(\mathfrak{b}_n \leqslant 0) = H(\tau_n^{\mathfrak{b}}; n, \phi_{\mathcal{S}})$$

$$\tau_n^{\mathfrak{b}} = -\frac{a_{\mathcal{S},0}\theta_{\mathcal{W}} - (n + a_{\mathcal{W},0})\theta_{\mathcal{S}} - (a_n^2 + a_n)n - (a_n^2 + a_n)a_{\mathcal{W},0} + a_{\mathcal{S},0}a_n^2 + a_{\mathcal{S},0}a_n}{\theta_{\mathcal{W}} + \theta_{\mathcal{S}} + 2a_n(a_n + 1)}$$

whereas the probability to select self-employment is $\mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{S}) = 1 - H(\tau_n^{\mathfrak{b}}; n, \phi_{\mathcal{S}})$.

Proof. See Appendix A.1. ∎

Decisions, respectively strategies, are evaluated according to a simple criterion: the number of successes they produce. A straightforward evaluation criterion is the expected number of successes given by

$$C \equiv \mathbb{E}\left[\sum_{n=1}^{N} \Omega(\mathfrak{d}_n) | a_{\mathcal{S},0}, a_{\mathcal{W},0}\right] \tag{10}$$

where $\Omega(\mathfrak{d}_n) \in \{0, 1\}$ is the reward given choice $\mathfrak{d}_n$. C only evaluates individual decision performance, abstracting from welfare effects and other non-individual criteria.

We establish the following property for the behavioral patterns of unbiased utility:

Lemma 2. *Behaving according to relative unbiased utility maximizes C such that unbiased utility is an optimal strategy given $(a_{\mathcal{S},0}, a_{\mathcal{W},0})$. By implication, behavior induced by relative unbiased utility is optimal.*

Proof. See Appendix A.2. ∎

For biased utility, we obtain the following result:

Proposition 1. *In general, behaving according to relative biased utility does not maximize C such that behavior is not optimal. Wage work will be selected with a higher than optimal probability if the ambiguity aversion associated with self-employment is higher than the aversion associated with wage work or $\theta_{\mathcal{S}} > \theta_{\mathcal{W}}$, where $\theta_{\mathcal{S}}$ is sufficiently large. The same applies to self-employment that is selected with a higher than optimal probability if the ambiguity aversion associated with wage work is higher than the ambiguity aversion associated with self-employment or $\theta_{\mathcal{W}} > \theta_{\mathcal{S}}$, where $\theta_{\mathcal{W}}$ is sufficiently large.*

Proof. See Appendix A.3. ∎

In this section, we demonstrate that the individually optimal strategy, maximizing the expected sum of occupational successes, is to always select the option with the highest expected success probability. The optimal strategy prescribes to exclusively concentrate on expected successes and to ignore potential errors in form of deviations of the true success probability from the expected value.

However, individuals making occupational choices might not be able to fully ignore errors, where the possibility of errors represents ambiguity. Therefore, we introduce a second decision strategy that accounts for ambiguity and, more specifically, ambiguity aversion, while also allowing for ambiguity aversion to differ across occupational choices. The introduction of ambiguity aversion reveals that there might be a bias for or against a particular occupation if ambiguity aversion is asymmetric across occupational options. This particular bias—the difference in choice probabilities between the optimal strategy and a strategy accounting for ambiguity aversion—is necessarily produced by a sufficiently high level of asymmetric ambiguity aversion in our model but cannot be tested with the data available to use.

4 Impact of media on decision patterns

Does media change choice probabilities (behavior)? Moreover, when does the consumption of articles and reports favoring entrepreneurship by ambiguity-averse individuals decrease deviations from optimal behavior, thereby improving decisions? To answers both questions, we analyze the impact of media and illustrate the model's mechanism by depicting (potentially positive) effects of media on occupational choices. We also derive two predictions that can be empirically tested.

4.1 Media as an informational intervention

Media articles and reports are denoted by $\mathfrak{m} \in \mathbb{N}$. $\mathfrak{m}$ is an informational intervention that does not affect probabilities to succeed. One of the simplest ways to formalize such an informational shock in favor of self-employment is to assume that in period $n = 0$ individuals are shown $\mathfrak{m} > 0$ additional successes in self-employment. Given media, instead of prior information $a_{\mathcal{S},0}$, individuals base their decisions on $\mathring{a}_{\mathcal{S},0} = a_{\mathcal{S},0} + \mathfrak{m}$, where $\mathring{a}_{\mathcal{S},0} > a_{\mathcal{S},0}$, while prior information about wage work is not directly affected. Media intensity is measured by the size of $\mathfrak{m}$, i.e., an increase in $\mathfrak{m}$ is interpreted as an increase in intensity.

Using the definition of media introduced above, media effects with respect to behavior are as follows.

Proposition 2. *Let* $\mathring{\mathfrak{d}}_n$ *denote a choice affected by media, whereas the choice without media impact is* $\mathfrak{d}_n$. *Given sufficient media intensity* $\mathfrak{m} > 0$ *and the two decision rules established (viz., unbiased and biased utility), media increases the probability to select self-employment and decreases the probability of wage work such that* $\mathbb{P}(\mathring{\mathfrak{d}}^{\mathfrak{u}}_n = \mathcal{S}) > \mathbb{P}(\mathfrak{d}^{\mathfrak{u}}_n = \mathcal{S})$ *and* $\mathbb{P}(\mathring{\mathfrak{d}}^{\mathfrak{u}}_n = \mathcal{W}) < \mathbb{P}(\mathfrak{d}^{\mathfrak{u}}_n = \mathcal{W})$, *and* $\mathbb{P}(\mathring{\mathfrak{d}}^{\mathfrak{b}}_n = \mathcal{S}) > \mathbb{P}(\mathfrak{d}^{\mathfrak{b}}_n = \mathcal{S})$ *and* $\mathbb{P}(\mathring{\mathfrak{d}}^{\mathfrak{b}}_n = \mathcal{W}) < \mathbb{P}(\mathfrak{d}^{\mathfrak{b}}_n = \mathcal{W})$ *for all n, where the effect requires* $\theta_{\mathcal{S}} \geqslant \theta_{\mathcal{W}}$ *in case of biased utility.*

Proof. See Appendix A.4. ■

The number of settings where positive media articles and reports about entrepreneurship might, theoretically, have a positive effect on occupational choices is restricted. For instance, if there is already a bias for self-employment, as might happen when $\theta_{\mathcal{W}} > \theta_{\mathcal{S}}$ (see Proposition 1), attempting to increase the number of self-employed is unnecessary. Yet, there is one setting where media does have normatively positive effects.

Proposition 3. *Assume that ambiguity aversion in self-employment is higher than ambiguity aversion in wage work, $\theta_{\mathcal{S}} > \theta_{\mathcal{W}}$, such that individual decisions are biased against self-employment. In such a setting, there always exists a level of media intensity such that the bias against self-employment is reduced. However, too intensive media effects might also create a bias for self-employment.*

Proof. See Appendix A.5. ■

To build intuition on the model's mechanism, consider a simple numerical example demonstrating how media influences (and improves) decisions.

Example 2. Assume that we could observe a sufficiently high number of alternative decision histories, allowing us to evaluate choice distributions, of an individual who retires after 50 periods. The individual's true probability to succeed in wage work is 20% and the probability to succeed in self-employment is 80%. Both probabilities are unknown to the individual making decisions. Furthermore, before her own career, the individual could observe the careers of two relatives. One relative was successful in self-employment over 5 periods, whereas the other relative was successful in wage work over 10 periods. Consequently, initial information suggests an expected success probability of 33% in self-employment and 67% in wage work.

Assume that the individual is not particularly ambiguity-averse but that ambiguity aversion with respect to self-employment is substantially higher than with respect to wage work such that $\theta_{\mathcal{W}} = 100$ and $\theta_{\mathcal{S}} = 5\theta_{\mathcal{W}}$.

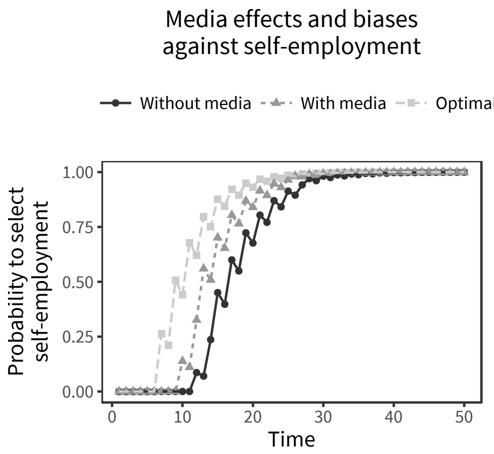
The individual selects self-employment with a lower than optimal probability, which is depicted in Figure 2a, and wage work with a higher than optimal probability, which is depicted in Figure 2b, because her ambiguity preferences bias her toward wage work. Now, assume that the individual watched TV reports about successful entrepreneurs. She decided that two reports ($\mathfrak{m} = 2$) were trustworthy. Hence, media only slightly changes the expected probability to succeed in self-employment, which increases by about 8 percentage points, and the probability to succeed in wage work, which decreases by 8 percentage points. Yet, even the small change increases the probability to become self-employed, respectively reduces the probability to select wage work, as shown in Figures 2a and 2b. As a consequence of incorporating information from the media into her beliefs, the individual becomes more successful, as demonstrated in Figure 2c, where the sum of successes given media effects clearly dominates successes without media effects.

Thus, in our model, positive media articles and reports about entrepreneurship increase the probability to select self-employment, whereas the probability to select wage work decreases. Furthermore, if ambiguity aversion in self-employment is sufficiently higher than ambiguity aversion associated with wage work, there will be a bias against self-employment. Positive media articles and reports about entrepreneurship can help reducing biases. Reducing a bias against self-employment involves an informational "push" towards self-employment.

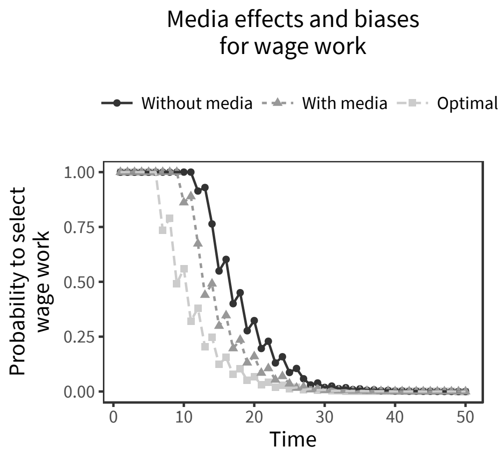
4.2 Predictions

The theoretical model allows us to formulate the following two predictions, on the basis of Proposition 2, with respect to marginal effects of media on occupational choices:

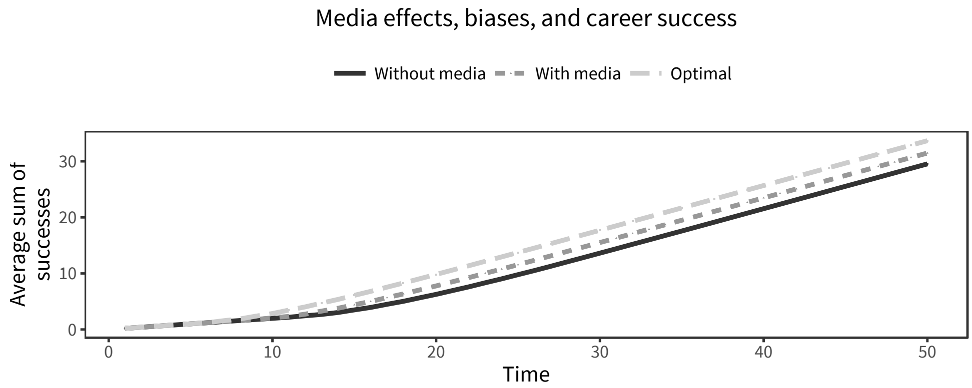
Hypothesis 1. In the wake of consuming media articles and reports with positive attitudes toward entrepreneurship, the probability of self-employment increases.



(a) Media and bias against self-employment



(b) Media and bias for wage work



(c) Media and career success

FIGURE 2. Media corrects a bias against self-employment

HYPOTHESIS 2. The consumption of media articles and reports about successful entrepreneurs reduces the probability of selecting wage work.

5 Empirical evidence on media effects

Having discussed how media affects occupational choices in theory, in this section we empirically investigate whether positive media articles and reports about entrepreneurship influence choices. Therefore, we formulate an empirical strategy to test whether behavior predicted by our theoretical model is consistent with actual occupational choice behavior.

We use two data sets and, thus, construct two empirical models. The first data set is a micro panel based on US data (the Integrated Health Interview Series). Media effects in the micro panel are identified with a heteroskedastic IV probit approach. The second data set is a macro panel of 38 countries, where media effects are identified with a linear IV regression model. Additionally, to ensure the robustness of our results, we also use different media variables in the micro and macro panel.

To induce an exogenous variation in positive media articles and reports about entrepreneurship, we use the occurrence of natural disasters in other regions and countries as an instrument. It is unlikely that this particular variable is connected to factors driving occupational choice—

for instance, it is rather unlikely that a natural disaster in New Zealand has a direct effect on choice probabilities in the Midwestern United States—but the consumption of media articles and reports is influenced by the occurrence of natural disasters.[12]

5.1 A minimalistic model of media consumption

Before presenting data and regression approaches, we introduce a simple model of media consumption linking individual consumption of articles and reports about entrepreneurship to natural disasters, substantiating the first stage of our regressions.

Let vector $\mathbf{m} = [\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_k, \mathfrak{m}, \mathcal{U}]^\top$ denote media consumption consisting of reports not related to entrepreneurship, $\mathcal{R}_1, \ldots, \mathcal{R}_k$, positive reports about entrepreneurship, $\mathfrak{m}$, and urgent news, $\mathcal{U}$. Daily consumption time is restricted to 24 hours. Hence, we can safely assume that $\mathbf{1}^\top\mathbf{m} = \bar{m}$, where the consumption limit $\bar{m} > 0$ is fixed and $\mathbf{1}$ is a vector of ones with $k+2$ elements. Without loss of generality, assume that the only urgent news are news about disasters. Disasters induce a variation in $\mathcal{U}$ such that $\mathbb{V}[\mathcal{U}] > 0$. If $\mathbb{V}[\mathcal{U}] > 0$, we also have $\mathbb{V}[\bar{m} - \mathcal{U}] > 0$. Using $\mathbf{1}^\top\mathbf{m} = \bar{m}$, we obtain

$$\mathbb{V}[\bar{m} - \mathcal{U}] = \mathbb{V}\left[\sum_{l=1}^{k} \mathcal{R}_l\right] + \mathbb{V}[\mathfrak{m}] + 2\mathrm{Cov}\left[\sum_{l=1}^{k} \mathcal{R}_l, \mathfrak{m}\right] \tag{11}$$

Thus, if the variation induced by natural disasters is not completely absorbed by articles and reports not related to entrepreneurship and consumption is fixed at some level, two rather weak conditions, disasters will induce a variation in the consumption of positive articles and reports about entrepreneurship. This relation can be tested—by testing for instrument strength.

Whether disasters increase or decrease the consumption of articles and reports about entrepreneurship depends on the correlation between non-entrepreneurship-related news and stories about entrepreneurship. There might be a compensation effect, bad news (natural disasters) are compensated by reading success stories about entrepreneurs, or a crowding-out effect, individuals concentrate on bad news and reduce the consumption of stories about entrepreneurs. Our results (first-stage regressions) provide evidence for crowding-out effects.

5.2 Micro panel

5.2.1 Data description

Except for the media variable and its instrument, our micro panel is based on data from the Integrated Health Interview Series (IHIS; Minnesota Population Center and State Health Access Data Assistance Center 2016), which is in turn based on the National Health Interview Survey (NHIS). NHIS is an annual survey that has been conducted since 1957. NHIS mostly provides information on health but the survey also provides data on occupational choice and variables important for the choice such as previous income, work experience, education, access to finance, etc.[13] We only consider adults (18–65 years old) who are employed (either wage workers or self-employed) in the period 2004–2015.[14] Observations are either available at the individual

[12]There are results showing that physical destruction of infrastructure might foster entrepreneurial activities. However, even in case of the high-impact September 11 attacks on the World Trade Center entrepreneurial rebounding was restricted to the area of Manhattan (Paruchuri & Ingram 2012).
[13]For further information on the panel, see https://www.cdc.gov/nchs/nhis/.
[14]Google Trends data is available starting 2004.

level or at the level of US regions, as used by the United States Census Bureau, which are: Northeast, North Central (Midwest), South, and West. The panel is a repeated cross section. Given our restrictions on age and occupational status, 10,851 observations are available. Further information on the micro panel is provided in Appendix B.1. The variables used in the empirical analysis are as follows.

Occupational status
The dependent variable is binary (1 if an individual has a certain occupational status and zero else). The most common occupational status is wage work, which is shown in Figure B.2 (Appendix B.1). There are no striking differences in occupational shares between regions but the North Central (Midwest) region tends to have a smaller self-employment share than other regions.

First media variable: Consumption of articles about famous entrepreneurs
The consumption of entrepreneurial success stories is approximated by the regional frequency of the search item 'famous entrepreneurs' in Google. Data is provided by the Google Trends tool.[15] The tool provides results at the US state level, which are aggregated to obtain searches at the region level. Since results are always measured relative to the state with the most searches (which is normalized to 100), only effect directions can be identified.

The instrument: Number of natural disasters
The consumption of articles about famous entrepreneurs might be endogenous. Therefore, we instrument it by the number of natural disasters in non-US regions, as natural disasters are exogenous to occupational choice but are usually covered in media reports, thus, affecting the consumption of articles about entrepreneurs (see Section 5.1). Data on natural disasters is collected by the Centre for Research on the Epidemiology of Disasters (Guha-Sapir, Below & Hoyois 2016). We only consider natural or "complex"[16] disasters, while specifically excluding technological disasters, as the latter type is caused by human action and is less likely to be exogenous. Regions with disasters are assigned based on geographical and cultural proximity but we avoid assigning a region that is too close to the US region. The Northeast region is assigned disasters in Mexico; the North Central region is assigned disasters in Australia and New Zealand; the South region is assigned disasters in South America; finally, the West region is assigned disasters in Western Europe.

We include several major determinants of occupational choice identified in the previous literature, including a number of demographic characteristics, capital income, education levels, work experience, physical and mental health, as well as personality.

Demography
Demographic controls include age, gender, whether the individual was born in the United States, and ethnicity.

Income, education, and work
We also control for earnings during the previous year; whether the individual usually works full time; educational attainment, ranging from "never attended school" to "obtained a doctoral degree;" and years on main or longest, or last job. Furthermore, we control for whether the individual received public assistance or food stamps in the previous year, and if the individual

[15] Available under www.google.com/trends/.
[16] A complex disaster includes famines for which drought was not the main cause.

has access to the financial market, approximated by whether the individual earned dividends from stocks or mutual funds in the previous year.

Physical health

Health is controlled by a general health variable (increase indicates decreasing health) and more specifically by whether the individual has any activity limitations.

Mental health and personality

We also control for mental health by including a set of variables capturing individuals' answers to the question of whether everything felt like an effort in the past 30 days, whether feelings interfered with life; and how often the individual felt hopeless, nervous, restless, sad, or worthless. Note that mental health also partially captures personality traits, as traits are linked to the probability of depression and anxiety (Klein, Kotov & Bufferd 2011).

5.2.2 Identifying media effects in the micro panel

In the micro panel based on US data, we have observations at the level of individuals indexed by k and at the level of US regions indexed by r. At the individual level, the "panel" is a repeated cross section such that individuals (and their number) change from period to period, where time is indexed by n. The dependent variable is dichotomous. An individual k from region r in period n can be self-employed, $d_{k,r,n,\mathcal{S}} = 1$, or not self-employed, $d_{k,r,n,\mathcal{S}} = 0$. Furthermore, an individual can be a wage worker, $d_{k,r,n,\mathcal{W}} = 1$, or not a wage worker, $d_{k,r,n,\mathcal{W}} = 0$. As we only consider individuals who are employed, we must have $d_{k,r,n,\mathcal{W}} = 1$ if $d_{k,r,n,\mathcal{S}} = 0$ and $d_{k,r,n,\mathcal{S}} = 1$ if $d_{k,r,n,\mathcal{W}} = 0$.

To analyze media effects, we use a probit model. Let $d_{k,r,n,i} = \mathbb{1}\{d^*_{k,r,n,i} > 0\}$ for $i \in \mathbb{O}$ where $d^*_{k,r,n,i}$ is an unobserved latent variable. The latent variable is modeled as

$$d^*_{k,r,n,i} = \mathsf{e}_{r,i} + \kappa_{1,i} M^{[1]}_{r,n} + \boldsymbol{\rho}^\top_{1,i} \mathbf{x}^{[1]}_{k,r,n} + \nu^{[1]}_{k,r,n,i} \tag{12}$$

where $\mathsf{e}_{r,i}$ is an option-specific fixed region effect and $\mathbf{x}^{[1]}_{k,r,n}$ are individual- and region-specific controls. $\kappa_{1,i}$ is the reaction of the latent variable and, thus, the individual choice variable, to the regional consumption of positive media articles about entrepreneurs $M^{[1]}_{r,n}$.[17] $M^{[1]}_{r,n}$ is constructed on the basis of Google Trends data revealing information on the dynamics of the search item 'famous entrepreneurs.' Unfortunately, $\kappa_{1,i}$ does not allow for the identification of effects sizes, due to the construction of the media variable, but effect directions can be easily identified.

The error term is likely heteroskedastic. For instance, there is a gender gap in entrepreneurship (Wagner 2007). If women react differently than men to incentives to become self-employed, the variance of the error cannot be equal across all individuals. However, even though the choice model is normalized, heteroskedasticity results in biased parameter estimates in a probit model (Yatchew & Griliches 1985), which is, for instance, not the case in a linear model, where coefficients are still unbiased under heteroskedasticity. A straightforward approach to account for heteroskedasticity-related issues is to explicitly model its determinants (Alvarez & Brehm 1995) by including a subset of covariates in the error variance specification. Therefore, we assume that

$$\nu^{[1]}_{k,r,n,i} \sim \text{Normal}(0, \exp\{\mathbf{v}^\top \mathbf{z}_{k,r,n}\}) \tag{13}$$

[17] Note that by properties of the dependent variable, we must have $\kappa_{1,\mathcal{S}} = -\kappa_{1,\mathcal{W}}$ and $\boldsymbol{\rho}_{1,\mathcal{S}} = -\boldsymbol{\rho}_{1,\mathcal{W}}$.

where determinants of heteroskedasticity in $\mathbf{z}_{k,r,n}$ and covariates in $\mathbf{x}^{[1]}_{k,r,n}$ can partially overlap.

To account for a potential endogeneity of media consumption, we instrument it by disasters in other (non-US) regions $D^{[1]}_{r,n}$ (the construction of the instrument is explained in Section 5.2.1), yielding the following first stage:

$$M^{[1]}_{r,n} = \breve{e}_r + \varkappa_1 D^{[1]}_{r,n} + \breve{\boldsymbol{\rho}}_1^{\top} \mathbf{x}^{[1]}_{k,r,n} + \breve{\nu}^{[1]}_{k,r,n} \tag{14}$$

where $\breve{e}_r$ is a region fixed effect.

5.2.3 *Estimation and results of micro panel models*

Micro panel models are estimated by maximum likelihood. As choice incentives might vary between genders (Wagner 2007), we include gender in the variance model, in Equation (13). In addition, we also use the following covariates to model the variance: region fixed effects, age, ethnicity, health, and education. The variance determinants were selected from a larger set on the basis of statistical significance and plausibility. For instance, it is plausible that health matters for the reaction to incentives to become self-employed or wage worker, as does the age of the individual (Caliendo, Fossen & Kritikos 2014).

Table 1 presents first-stage estimation results. Instrument strength does not pose a problem (the F-statistic is 41.382). Disasters are negatively correlated to the consumption of articles about famous entrepreneurs, corresponding to a crowding-out effect, which is in line with the minimalistic model of media consumption, constructed in Section 5.1. The crowding out is consistent with previous research. There is, for example, evidence that humans are predisposed to focus on negative information, because the costs of ignoring negative information outweigh the benefits of positive information (Soroka & McAdams 2015). Such a negativity bias is a reasonable heuristic if costs and benefits from different types of information are asymmetric.

Table 2 shows results generated by heteroskedastic IV probit.[18] We observe that women are less likely to become self-employed, which is consistent with previous results (Cowling & Taylor 2001; Wagner 2007; Caliendo et al. 2014). Being female also has a significant effect on the variance of choices. Furthermore, in line with previous findings, work experience (years on job) and receiving dividends (access to the financial market) both increase the probability of self-employment (Blanchflower & Oswald 1998; Gompers, Lerner & Scharfstein 2005; Elfenbein, Hamilton & Zenger 2010), while the effect of age follows an inverse u-shaped relationship (Caliendo et al. 2014). Thus, the coefficients of non-media variables confirm earlier findings.

Turning now to the influence of our first media variable, the consumption of articles about famous entrepreneurs, we can see that the consumption of articles about famous entrepreneurs significantly increases the probability of selecting self-employment and reduces the probability of selecting wage work. The effects in Table 2 support Hypothesis 1 and 2.

5.3 Macro panel

5.3.1 *Data description*

To check whether results are robust, we also use, in addition to the micro data model, an empirical model based on an unbalanced country-level macro panel. Effects on the probability of self-employment and wage work are estimated on the basis of 38 countries. In sum, there are

[18] Estimation results indicate an endogeneity issue: Results with IV and without IV, given in Table C.1 (Appendix C), substantially differ.

TABLE 1. First stage of micro panel model, where dependent variable is media attention

Variable	Coefficient	SE
Number of natural disasters	–0.008***	(0.001)
Age	0.005	(0.003)
Age^2	–0.000	(0.000)
Female†	–0.006	(0.011)
Born in US†	0.049***	(0.017)
Non-white†	–0.042***	(0.013)
Full-time work†	0.038***	(0.013)
Earnings	–0.006**	(0.003)
Got dividends†	0.098***	(0.019)
Got food stamps†	0.501***	(0.030)
Got welfare†	–0.174***	(0.049)
Education	–0.001	(0.002)
Years on job	0.001	(0.001)
Limitations†	–0.027	(0.016)
Health‡	–0.001	(0.006)
Effort	–0.013***	(0.005)
Feelings interfered with life	–0.002	(0.007)
Hopeless	–0.016**	(0.008)
Nervous	–0.014***	(0.005)
Restless	–0.003	(0.005)
Sad	0.016**	(0.006)
Worthless	0.008	(0.008)
North Central†	0.163***	(0.031)
Northeast†	–0.709***	(0.036)
West†	–1.271***	(0.030)
Constant	1.581***	(0.075)
	10,851 obs.; R^2 = 0.550	

Notes: F-statistic for instrument weakness with heteroskedasticity-robust errors: 41.382; †dummy variable; ‡increase indicates more health problems; ***significant at the 1%-level; **significant at the 5%-level; standard errors in parentheses are heteroskedasticity-consistent.

170 joint observations. The panel is fairly representative, as both developed and developing countries are included.[19] We use annual country-level data from four different sources: the Global Entrepreneurship Monitor, the World Bank, Transparency International, and the Centre for Research on the Epidemiology of Disasters. We only consider the 2003–2012 time period. The minimum of observed periods is 3 and the maximum is 10. Appendix B.2 shows data characteristics. The following variables are used in our regressions:

Shares of self-employed and wage workers

The dependent variables of our regression models are relative choice frequencies, or empirical probabilities. We approximate relative choice frequencies by the share of wage workers and self-employed provided by the World Bank. In Figure B.3 (Appendix B.2) it is shown that there is a substantial variation in choice frequencies across countries. For instance, the maximum country

[19]The following countries are included: Argentina, Australia, Belgium, Brazil, Canada, Chile, Colombia, Croatia, Ecuador, Greece, Hong Kong, Hungary, Iran, Ireland, Israel, Italy, Jamaica, Japan, South Korea, Latvia, Malaysia, Mexico, the Netherlands, New Zealand, Norway, Peru, Poland, Romania, Russia, Serbia, Singapore, Slovenia, Sweden, Switzerland, the UK, the USA, Uruguay, and Venezuela.

TABLE 2. IV probit estimates of marginal effects in micro panel model, where dependent variable is choice dummy

	Self-employment		Wage work	
Variable	Coefficient	SE	Coefficient	SE
Consumption of articles about famous entrepreneurs§	0.611***	(0.196)	–0.611***	(0.196)
Age	0.178***	(0.034)	–0.178***	(0.034)
Age^2	–0.002***	(0.000)	0.002***	(0.000)
Female†	–1.912***	(0.525)	1.912***	(0.525)
Born in US†	–0.191	(0.100)	0.191	(0.100)
Non-white†	–0.234	(0.158)	0.234	(0.158)
Full-time work†	–0.405***	(0.108)	0.405***	(0.108)
Earnings	0.004	(0.013)	–0.004	(0.013)
Got dividends†	0.237**	(0.119)	–0.237**	(0.119)
Got food stamps†	–0.369**	(0.177)	0.369**	(0.177)
Got welfare†	0.167	(0.251)	–0.167	(0.251)
Education	0.019	(0.017)	–0.019	(0.017)
Years on job	0.073***	(0.016)	–0.073***	(0.016)
Limitations†	0.060	(0.091)	–0.060	(0.091)
Health‡	0.012	(0.059)	–0.012	(0.059)
Effort	0.034	(0.028)	–0.034	(0.028)
Feelings interfered with life	–0.045	(0.041)	0.045	(0.041)
Hopeless	–0.071	(0.047)	0.071	(0.047)
Nervous	0.018	(0.031)	–0.018	(0.031)
Restless	0.071**	(0.031)	–0.071**	(0.031)
Sad	–0.035	(0.038)	0.035	(0.038)
Worthless	–0.070	(0.049)	0.070	(0.049)
North Central†	–0.456**	(0.197)	0.456**	(0.197)
Northeast†	0.302	(0.198)	–0.302	(0.198)
West†	0.610**	(0.260)	–0.610**	(0.260)
Constant	–5.656***	(0.763)	5.656***	(0.763)
Variable: Variance model		Coefficient	SE	
North Central†		–0.042	(0.070)	
Northeast†		–0.100	(0.083)	
West†		0.057	(0.073)	
Age		0.013***	(0.003)	
Female†		0.584***	(0.109)	
Non-white†		–0.083	(0.067)	
Health‡		–0.052**	(0.026)	
Education		–0.003	(0.008)	
		10,851 obs.		

Notes: §Normalized media consumption is instrumented by number of natural disasters in other regions and countries; †dummy variable; ‡increase indicates more health problems; ***significant at the 1%-level; **significant at the 5%-level; standard errors in parentheses are heteroskedasticity-consistent.

average of the probability of self-employment is 52% (in Peru), while the minimum level is 7% (in the US).

Second media variable: Media attention for entrepreneurship

The media variable in the macro panel differs from the one in the micro model. Instead of deriving it from Google Trends data, we approximate media intensity by "media attention for entrepreneurship," surveyed by the Global Entrepreneurship Monitor (GEM). Media attention is measured by the percentage of the population aged 18–64 who report that in their country there are frequent media reports about successful new businesses. The advantage of the indicator provided by the GEM compared to other indicators is that it measures perceptions so that we can be sure that decision makers are aware of positive reports about entrepreneurship.

The instrument: Number of natural disasters

As in the micro model, we instrument the media variable by the number of natural disasters in other countries. In the macro panel, countries are paired randomly and the combination yielding the strongest instrument is selected, as described in Section 5.3.2. Note that we have more observations of natural disasters than of media attention for entrepreneurship such that the pool of countries with disasters is larger than 38.

Fear of entrepreneurial failure

Countries differ with respect to their attitudes towards entrepreneurial failure. As noted by the *Economist*:

> If you start a company in London or Paris and go bust, you have just ruined your future; do it in Silicon Valley and you have simply completed your entrepreneurial training. (The Economist 1997, p. 17)

Being afraid to fail, and the associated stigma, can prevent an individual from becoming self-employed. Hence, we control for country-specific attitudes towards failure by including the percentage of the population aged 18–64 perceiving good opportunities for business who indicate that fear of failure would prevent them from setting up a business. The fear of failure rate is provided by the GEM.

Ease of doing business

Annual indicators of ease of doing business measuring a country's regulatory environment[20] are provided by the World Bank Group (Doing Business project). The higher the indicator value is, the easier is doing business. In our sample, ease of doing business mostly reflects the difference between developed and developing countries—doing business tends to be easier in developed countries. The difference in economic development is important for occupational choice as, compared to developed countries, developing countries are exposed to higher unemployment levels, have lower levels of wage work, and higher levels of self-employment (see, e.g., Chen & Doane 2008; Gindling & Newhouse 2012). Unfortunately, using annual indicators would greatly reduce the number of available observations. Therefore, we, first, take country-specific averages and, then, construct two groups based on these country averages with k-means clustering: a group of countries where doing business is relatively easy and a group where it is relatively difficult. This classification is assumed to hold for all periods 2003–2012, even if annual ease of doing business was not observed in some periods.[21]

[20] The regulatory environment includes components such as starting a business, dealing with construction permits, getting electricity, registering property, getting credit, protecting minority investors, paying taxes, trading across boarders, enforcing contracts, and resolving insolvency (see World Bank Group 2015).

[21] This assumption makes sense if relative ease of doing business is sufficiently stable over time. In Appendix B.3, we examine stability with available data and find a strong tendency of countries to remain in one group.

Corruption
The Corruption Perceptions Index, taking values on the interval [0, 10], is annually provided by Transparency International. The higher the index value is, the *less* corruption is perceived. We include a measure of corruption in our regressions because our data includes developing countries and previous research shows the relative importance of institutional constraints impeding development in developing economies (Goedhuys & Sleuwaegen 1999; Ardagna & Lusardi 2010; Quatraro & Vivarelli 2015). Furthermore, Anokhin & Schulze (2009) demonstrate that corruption hampers innovation and entrepreneurship.

Other controls
In addition to the aforementioned variables, we use the following controls that could also affect occupational choices: GDP (per capita), GDP growth, inflation, and the real interest rate. All four covariates are provided by the World Bank.

5.3.2 *Identifying media effects in the macro panel*

We have data for a set of countries indexed by j. Each country j is observed over some periods indexed by n. The number of observed periods is allowed to differ across countries.[22] With respect to the dependent variable, we observe two shares for each j and n. The first share, $p_{j,n,\mathcal{S}} \in (0,1)$, is the share of self-employed individuals (the empirical probability of the choice 'self-employment') in the working-age population. The second share, $p_{j,n,\mathcal{W}} \in (0,1)$, is the share of wage workers in the working-age population. We also refer to $p_{j,n,i}$ as the relative frequency of occupation $i \in \mathbb{O}$.

Let $\mathfrak{L}(p) \equiv \log(p[1-p]^{-1})$ denote the logit transformation function, where $p \in (0,1)$ is a proportion. The transformation maps a share on the real line. To model a relative choice frequency, we use the following linear model:

$$\mathfrak{L}(p_{j,n,i}) = \mathsf{c}_i + \kappa_{2,i} M^{[2]}_{j,n} + \boldsymbol{\rho}^{\top}_{2,i}\mathbf{x}^{[2]}_{j,n} + \nu^{[2]}_{j,n,i} \tag{15}$$

$\mathbf{x}^{[2]}_{j,n}$ are time- and country-specific covariates. c_i is an option-specific constant. $\kappa_{2,i}$ is the option-specific effect of media, i.e., the effect of most interest. $\exp(\kappa_{2,i})$ corresponds to the relative change in odds given a one unit increase in media attention, when all the remaining variables are held constant. To approximately examine effect sizes, we also use a linear probability model, where the left hand side of (15) is $p_{j,n,i}$.

Equation (15) is the second stage of our regression. As the media variable in (15) might be endogenous, $M^{[2]}_{j,n}$ and the error term $\nu^{[2]}_{j,n,i}$ may be correlated, we instrument media attention by the number of natural disasters, denoted by $D^{[2]}_{j,n}$. To ensure that the exclusion restriction holds, we only use disaster data from other countries.[23] The first stage is as follows:[24]

$$M^{[2]}_{j,n} = \breve{\mathsf{c}} + \varkappa_2 D^{[2]}_{j,n} + \breve{\boldsymbol{\rho}}^{\top}_2 \mathbf{x}^{[2]}_{j,n} + \breve{\nu}^{[2]}_{j,n} \tag{16}$$

To generate an instrument with sufficient strength, we use the following three-step approach:

[22] However, we require that $n \geqslant 3$ for all j such that the effects of time-variant variables can be distinguished from the impact of time-invariant covariates.
[23] $D^{[2]}_{j,n}$ captures disasters in a country assigned to j but different from j.
[24] We do not transform $M^{[2]}$—given our data, $M^{[2]} \in (0,1)$—as this would limit interpretations. However, our main results, the outcome of the test of the two central model predictions, does not depend on the transformation of $M^{[2]}$.

STEP 1. Each country i in the panel is randomly, without repetitions, assigned another country $\mathfrak{a}_i$ with disasters resulting in assignment matrix

$$\mathcal{A}_{\mathsf{r}} = \begin{bmatrix} \text{Country 1} & \mathfrak{a}_1 \\ \text{Country 2} & \mathfrak{a}_2 \\ \vdots & \vdots \\ \text{Country 38} & \mathfrak{a}_{38} \end{bmatrix}$$

The assignment procedure is repeated R times, such that we obtain the general assignment matrix

$$\mathbf{A} = \begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \cdots & \mathcal{A}_{\mathsf{R}-1} & \mathcal{A}_{\mathsf{R}} \end{bmatrix}$$

STEP 2. For each assignment $\mathcal{A}_{\mathsf{r}}$ in $\mathbf{A}$, a first-stage F-statistic, to assess instrument strength, is computed (Staiger & Stock 1997). Statistics account for heteroskedasticity or clustering at the country level.

STEP 3. The combination with the best F-statistic result given potential heteroskedasticity, conditional on sufficient instrument strength in case of errors clustering at the country level, is selected.

Instrument strength is considered as sufficient if the first-stage partial F-statistic is substantially larger than 10 (Staiger & Stock 1997; Stock & Yogo 2005). Besides the best instrument, the three-step approach will generate a number of country pairings with sufficient strength (a large F-statistic). These combinations can be used to test whether results depend on a particular combination of countries or are robust to using different country pairs.

5.3.3 *Estimation and results of macro panel models*

To estimate our macro panel models, we use two-stage least squares (with heteroskedasticity-robust standard errors and errors clustered at the country level). The best pairing of countries on the basis of 10,000 random assignments is given in Table C.3 (Appendix C). Table 3 shows first-stage results. According to first-stage F-statistics, which are both larger than 100, the instrument is sufficiently strong independent of whether we use heteroskedasticity-robust errors or cluster errors at the country level. As in the micro panel model, natural disasters are negatively correlated to the media variable: There is a crowding-out effect. An additional natural disaster is associated with a reduction in media attention for entrepreneurship of 1.4 percentage points.

Table 4 presents results generated by IV regressions, with Table 3 as first stage.[25] As positive media reports about entrepreneurship increase the probability of self-employment, we find support for Hypothesis 1. Table 4 also provides support for Hypothesis 2: Media reports about entrepreneurial success reduce the probability of wage work. The effects are driven by differences *between* countries, as after the inclusion of country fixed effects (not presented here[26]) media effects become insignificant.

The exponential of the coefficient of media attention can be interpreted as an effect on odds[27]

[25]Two-stage least squares results significantly differ from OLS results, given in Table C.2 (Appendix C). OLS tends to underestimate effects.
[26]Results available from the authors on request.
[27]The odds of occupation $i \in \mathbb{O}$ are $p_i(1 - p_i)^{-1}$.

TABLE 3. First stage of macro panel model, where dependent variable is media attention

Variable	Coefficient	SE
Number of natural disasters	–1.442***[***]	(0.155)
Doing business is relatively easy†	5.819	(3.174)
Fear of entrepreneurial failure	–29.984**	(12.160)
Inflation	0.346	(0.211)
GDP	0.000	(0.000)
GDP growth	0.966***[***]	(0.223)
Real interest rate	0.523***[***]	(0.076)
Lack of corruption‡	0.895	(0.725)
Constant	64.653***[***]	(6.425)
	170 obs.; R^2 = 0.418	

Notes: F-statistic for instrument weakness with heteroskedasticity-robust errors: 86.929; F-statistic for instrument weakness with errors clustered at country level: 161.737; †dummy is 1 if yes and zero else; ‡increase indicates less corruption; ***significant at the 1%-level; **significant at the 5%-level; [***]significant at the 1%-level with country-level clustering; [**]significant at the 5%-level with country-level clustering; standard errors in parentheses are heteroskedasticity-consistent.

TABLE 4. IV estimates of marginal effects in macro panel model, where dependent variable is transformed choice share

	Self-employment		Wage work	
Variable	Coefficient	SE	Coefficient	SE
Media attention for entrepreneurship§	0.028***[***]	(0.004)	–0.021***[***]	(0.003)
Doing business is relatively easy†	–0.678***[**]	(0.158)	0.473***	(0.121)
Fear of entrepreneurial failure	–0.014	(0.667)	0.163	(0.532)
Inflation	–0.023	(0.015)	0.017	(0.012)
GDP	0.000***	(0.000)	0.000***[**]	(0.000)
GDP growth	–0.008	(0.014)	0.016	(0.011)
Real interest rate	–0.005	(0.004)	0.001	(0.003)
Lack of corruption‡	–0.047	(0.033)	0.074***	(0.023)
Constant	–2.025***[***]	(0.392)	0.958	(0.301)
		170 obs.		

Notes: §Media attention is instrumented by number of natural disasters in other countries; †dummy is 1 if yes and zero else; ‡increase indicates less corruption; ***significant at the 1%-level; **significant at the 5%-level; [***]significant at the 1%-level with country-level clustering; [**]significant at the 5%-level with country-level clustering; standard errors in parentheses are heteroskedasticity-consistent.

when all other variables are held constant. A one percentage point increase in positive media attention for entrepreneurship increases the odds of self-employment by 2.8% and decreases the odds of wage work by 2.1%.

5.3.4 *Robustness*

Assignment of countries

Using a different assignment of countries with natural disasters produces similar effects. In Figure 3, we show the estimated effects on self-employment and wage work (all significant at the 5%-level using errors clustered at the country level) of the 15 best unique assignments.

The minimum effect on self-employment is 0.01 (effect on odds: increase by 0.92%), the maximum is 0.04 (effect on odds: increase by 4.52%), whereas the average effect is 0.02 (effect on odds: increase by 2.51%). The minimum effect on wage work is –0.01 (effect on odds:

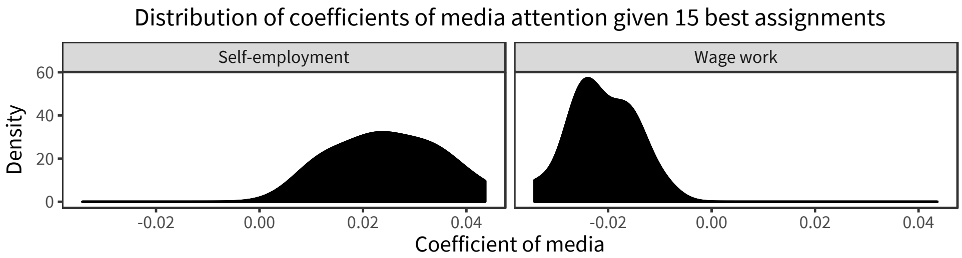


FIGURE 3. Using different country pairings, resulting in different instruments

decrease by 0.89%), the maximum is –0.03 (effect on odds: decrease by 3.40%), while the average effect is –0.02, which is the effect established with the best assignment.

Beta model as alternative to logit transformation

There is an open concern that our models might be misspecified because the logit transformation does not fully remove skewness from our dependent variables. For instance, the distribution of transformed wage work shares in Figure B.4 (Appendix B.2) is clearly skewed. Thus, our results might be mostly driven by modeling assumptions.

To reduce the danger of model misspecification (especially, the danger that results are driven by skewness), we model the original, non-transformed, shares with beta regressions.[28] The beta regression, proposed by Ferrari & Cribari-Neto (2004), accommodates skewness and heteroskedasticity, as values near zero and 1 have typically a smaller variance than other values in the $(0, 1)$ interval. We employ a two stage procedure. The first stage is (16), estimated in Table 3, while the second stage is the beta regression. We use estimated residuals from the first stage as an additional predictor in the second stage (Newey 1987; Terza, Basu & Rathouz 2008).[29] Second stage confidence intervals are bootstrapped with clustering at the country level in line with Efron (1987).

Beta regression results presented in Table 5 clearly support Hypothesis 1 and 2, as media attention for entrepreneurship significantly increases the average share of the self-employed and reduces the average share of wage workers. Thus, the micro and macro panel models, based on two different data sets, support Hypothesis 1 and 2, derived from our theoretical model of career choice under ambiguity.

5.3.5 *Effect sizes*

In the micro panel, effect sizes cannot be properly interpreted, because of the construction of the Google Trends variable. However, the macro panel allows for a simple interpretation. To approximate effect sizes, we estimate linear probability models, where Table 3 is the first stage.

Figure 4a shows effects of a 1 percentage point increase in media attention for entrepreneurship, including 95% confidence intervals (full results are in Table C.4 in Appendix C). The probability to select self-employment increases by 0.47 percentage points and the probability to select wage work decreases by 0.44 percentage points.

In Figure 4b, we compare the persuasion effect established by us—the 0.5 percentage point increase in the probability to select self-employment—to persuasion effects found in the liter-

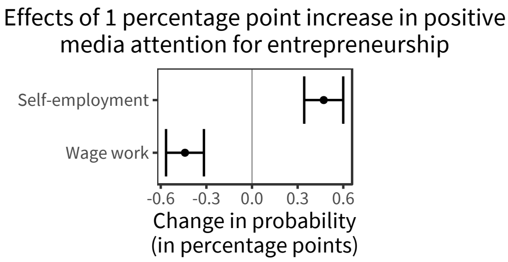
[28]In line with suggestions of Ferrari & Cribari-Neto (2004), the beta distribution is parameterized in terms of its mean and precision (a large precision corresponds to a small variance). A linear combination of predictors is linked to the mean by a logit link. Consequently, a positive estimated coefficient of a predictor can be interpreted as a positive effect on the average share and *vice versa*.

[29]Using predicted values from the first stage yields numerically very similar media effects.

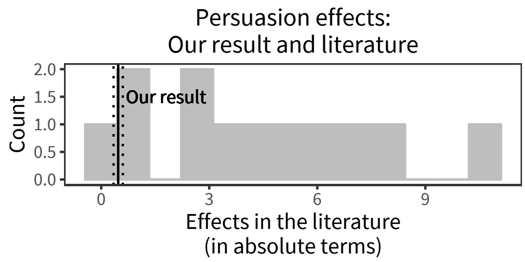
TABLE 5. IV beta estimates of marginal effects on original shares using macro panel, where dependent variable is original choice share

	Self-employment		Wage work	
Variable	Coefficient	95% CI	Coefficient	95% CI
Media attention for entrepreneurship	0.028**	0.02, 0.04	–0.021**	–0.03, –0.01
Residuals from first stage	–0.017**	–0.03, –0.01	0.013**	0.00, 0.02
Doing business is relatively easy†	–0.592**	–0.90, –0.28	0.433**	0.18, 0.68
Fear of entrepreneurial failure	–0.306	–1.62, 0.87	0.398	–0.54, 1.42
Inflation	–0.028**	–0.07, –0.00	0.019	–0.00, 0.05
GDP	0.000**	0.00, 0.00	0.000**	0.00, 0.00
GDP growth	–0.009	–0.05, 0.02	0.018	–0.00, 0.05
Real interest rate	–0.006	–0.01, 0.00	0.002	–0.01, 0.01
Lack of corruption‡	–0.057	–0.11, 0.00	0.077**	0.04, 0.12
Constant	–1.711**	–2.56, –0.98	0.801**	0.18, 1.41
Precision parameter	36.263**	27.89, 41.43	46.497**	34.59, 54.74
		170 obs.		

Notes: †Dummy is 1 if yes and zero else; ‡increase indicates less corruption; **significant at the 5%-level; confidence intervals (CI) are bootstrapped (2,000 replications) at the country level.



(a) Estimated effect sizes



(b) Persuasion

FIGURE 4. Effect sizes in macro panel according to linear probability models

ature[30] on media effects (viz., Ansolabehere & Iyengar 1995; Gerber & Green 2000; Green, Gerber & Nickerson 2003; Kull, Ramsay & Lewis 2003; Gentzkow & Shapiro 2004; Della Vigna & Kaplan 2007; Gerber, Karlan & Bergan 2009). Our result is consistent with previous findings but the effect size is rather small in comparison with other studies. However, the standard deviation of the media variable in our sample is approximately 15 percentage points so that even the small effect size leads to substantial effects given the variation of positive media attention for entrepreneurship. The effect of media on the probability of self-employment is comparable to the effect of watching the Fox News channel on the Republican vote share (a gain of 0.4 to 0.7 percentage points) found by Della Vigna & Kaplan (2007).

5.4 Limitations and further research

Our approach has several limitations. First of all, as already mentioned, we cannot directly test the theoretical model, but only test for consistency of derived hypotheses. Also, as we are not

[30] We use data on media effects compiled by Della Vigna & Kaplan (2007, Table IX). The persuasion effect is computed as the absolute difference in the outcome variable between treatment and control group. All outcome variables are shares so that effects are comparable to our results.

able to compute optimal choice probabilities in our empirical analysis, we cannot determine whether there are too many or too few choices of self-employment. Therefore, our empirical model is restricted to testing whether positive media reports about famous entrepreneurs influence choices in the expected direction.

Furthermore, we cannot identify voluntary unemployment in data. However, preliminary results derived from the theoretical model suggest that media reduces voluntary unemployment by reducing ambiguity about employment options. A reduction in voluntary unemployment caused by media could be seen as a positive effect since unemployment generates negative psychological effects (Paul & Moser 2009) and impairs the generation of valuable information about success probabilities leading to lower life-time earnings (Verbruggen, van Emmerik, Van Gils, Meng & de Grip 2015).

Last but not least, the levels of aggregation of our media variables are rather high (region and country levels). It would be more preferable to use media consumption at the individual or household level, and observe the household or individual over a sufficiently long period of time. This would allow to determine more directly whether media influences individuals in their occupational choices. Consequently, further research is necessary.

6 Concluding summary

There is no question that media wields significant power in modern societies. Media reports can reverse corporate governance violations (Dyck et al. 2008) or influence voting behavior (Della Vigna & Kaplan 2007). Our theoretical and empirical analysis adds a new effect to the literature: Media affects occupational choices.

In our theoretical model, we show under what conditions media influences occupational choices, and in which circumstances this influence is positive. We demonstrate that ambiguity-averse individuals might not make individually optimal choices, in the sense that they are not selecting those options that yield the highest expected success probabilities. Instead, they might exhibit a bias for, respectively against, some occupational option due to asymmetric ambiguity aversion. We show that sufficiently intensive positive media reports about entrepreneurs, transporting ambiguity-reducing information, increase the probability of selecting self-employment, while the probability of wage work is reduced. In case of asymmetric ambiguity preferences biased against self-employment, when ambiguity aversion related to self-employment is sufficiently higher than to wage work, media reduces a behavioral bias against self-employment.

Given micro-level data and country-level panel data, we test central predictions from our theoretical model, in particular to what extent media affects choice probabilities. For that reason, we estimate multiple instrumental variable regressions to determine the empirical effects of media. In line with our theoretical model, we establish that the consumption of positive media articles and reports about entrepreneurs significantly increases the probability of self-employment and significantly reduces the probability of wage work.

To conclude, media reports can foster self-employment, while reducing wage work, by providing information that changes individual beliefs. Informational shocks can, thus, have a significant impact on career choices. The established effects are sufficiently large to be of interest; allow for a causal interpretation; are based on observations from two data sets using two different empirical models; and are robust to model specification. However, our regression approaches, relying on repeated cross sections or aggregated data, cannot directly evaluate individual decision histories. Yet, our theoretical model can be used to generate further testable predictions with respect to individual short- and long-run decision behavior, opening up venues for further research.

References

Alvarez, R. & Brehm, J. (1995). American Ambivalence Towards Abortion Policy: Development of a Heteroskedastic Probit Model of Competing Values. *American Journal of Political Science*, *41*(4), 345–374.

Anokhin, S. & Schulze, W. (2009). Entrepreneurship, Innovation, and Corruption. *Journal of Business Venturing*, *24*(5), 465–476.

Ansolabehere, S. & Iyengar, S. (1995). *Going Negative: How Attack Ads Shrink and Polarize the Electorate*. New York: Free Press.

Antonovics, K. & Golan, L. (2012). Experimentation and Job Choice. *Journal of Labor Economics*, *30*(2), 333–366.

Ardagna, S. & Lusardi, A. (2010). Explaining International Differences in Entrepreneurship: The Role of Individual Characteristics and Regulatory Constraints. In J. Lerner & A. Schoar (Eds.), *International Differences in Entrepreneurship*. Univ. of Chicago Press.

Bergemann, D. & Hege, U. (2005). The Financing of Innovation. *RAND Journal of Economics*, *36*(4), 719–752.

Besley, T. & Burgess, R. (2002). The Political Economy of Government Responsiveness: Theory and Evidence from India. *Quarterly Journal of Economics*, *117*(4), 1415–1451.

Blanchflower, D. & Oswald, A. (1998). What Makes an Entrepreneur? *Journal of Labor Economics*, *16*(1), 26–60.

Bosma, N., Hessels, J., Schutjens, V., Van Praag, M., & Verheul, I. (2012). Entrepreneurship and Role Models. *Journal of Economic Psychology*, *33*, 410–424.

Bradt, R. N., Johnson, S. M., & Karlin, S. (1956). On Sequential Designs for Maximizing the Sum of n Observations. *Annals of Mathematical Statistics*, *27*(4), 1060–1074.

Caliendo, M., Fossen, F., & Kritikos, A. (2014). Personality Characteristics and The Decision to Become and Stay Self-Employed. *Small Business Economics*, *42*(4), 787–814.

Camerer, C. & Ho, T.-H. (1999). Experience-weighted Attraction Learning in Normal Form Games. *Econometrica*, *67*(4), 827–874.

Camerer, C. & Weber, M. (1992). Recent Developments in Modeling Preferences: Uncertainty and Ambiguity. *Journal of Risk and Uncertainty*, *5*(4), 325–370.

Campanale, C. (2010). Private Equity Returns in a Model of Entrepreneurial Choice with Learning. *B.E. Journal of Macroeconomics*, *10*(1), 1–37.

Chen, M. & Doane, D. (2008). *Informality in South Asia: A Review*. Cambridge, MA: WIEGO.

Cowling, M. & Taylor, M. (2001). Entrepreneurial Women and Men: Two Different Species? *Small Business Economics*, *16*(3), 167–175.

Della Vigna, S. & Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics*, *122*(3), 1187–1234.

Djankov, S., Qian, Y., Roland, G., & Zhuravskaya, E. (2006). Entrepreneurship in China and Russia Compared. *Journal of the European Economic Association*, *4*(2/3), 352–365.

Dyck, A., Moss, D., & Zingales, L. (2008). Media versus Special Interest. *NBER Working Paper*, (14360).

Dyck, A., Volchkova, N., & Zingales, L. (2008). The Corporate Governance Role of the Media: Evidence from Russia. *The Journal of Finance*, *63*(3), 1093–1135.

Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association*, *82*(397), 171–185.

Eisensee, T. & Strömberg, D. (2007). News Droughts, News Floods, and U.S. Disaster Relief. *The Quarterly Journal of Economics*, *122*(2), 693–728.

Elfenbein, D., Hamilton, B., & Zenger, T. (2010). The Small Firm Effect and the Entrepreneurial Spawning of Scientists and Engineers. *Management Science*, *56*(4), 659–681.

Ellsberg, D. (1961). Risk, Ambiguity and Savage Axioms. *Quarterly Journal of Economics*, *75*(4), 643–679.

Enikolopov, R., Petrova, M., & Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia. *American Economic Review*, *101*(7), 3253–3285.

Fellner, W. (1961). Distortion of Subjective Probabilities as a Reaction to Uncertainty. *Quarterly Journal of Economics*, *75*(4), 670–689.

Ferrari, S. & Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics*, *31*(7), 799–815.

Fifield, A. (2016). Start-ups Encounter Hiccups in the Land of the Rising Sun. *The Washington Post*. 23 June 2016.

Frisch, D. & Baron, J. (1988). Ambiguity and Rationality. *Journal of Behavioral Decision Making*, *1*(3), 149–157.

Gentzkow, M. (2006). Television and Voter Turnout. *Quarterly Journal of Economics*, *121*(3), 931–972.

Gentzkow, M. & Shapiro, J. (2004). Media, Education, and Anti-Americanism in the Muslim World. *Journal of Economic Perspectives*, *18*(3), 117–133.

Gerber, A. & Green, D. (2000). The Effects of Canvassing, Telephone Calls and Direct Mail on Voter Turnout. *American Political Science Review*, *94*(3), 653–663.

Gerber, A., Karlan, D., & Bergan, D. (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal*, *1*(2), 35–52.

Gindling, T. & Newhouse, D. (2012). Self-employment in the Developing World. *World Bank Policy Research Working Paper*, (6201).

Gittins, J., Glazebrook, K., & Weber, R. (2011). *Multi-armed Bandit Allocation Indices* (2. ed.). Chichester: Wiley.

Goedhuys, M. & Sleuwaegen, L. (1999). Barriers to Growth of Firms in Developing Countries, Evidence from Burundi. In D. Audretsch & A. Thurik (Eds.), *Innovation, Industry Evolution and Employment*. Cambridge Univ. Press.

Gompers, P., Lerner, J., & Scharfstein, D. (2005). Entrepreneurial Spawning: Public Corporations and the Genesis of New Ventures, 1986 to 1999. *Journal of Finance*, *60*(2), 577–614.

Green, D., Gerber, A., & Nickerson, D. (2003). Getting Out The Vote in Local Elections. *Journal of Politics*, *65*(4), 1083–1096.

Guha-Sapir, D., Below, R., & Hoyois, P. (2016). EM-DAT: The CRED/OFDA International Disaster Database. Université Catholique de Louvain. Brussels, Belgium. www.emdat.be.

Hintermaier, T. & Steinberger, T. (2002). Occupational Choice and the Private Equity Premium Puzzle. *Journal of Economic Dynamics and Control*, *29*(10), 1765–1783.

Holm, H., Opper, S., & Nee, V. (2013). Entrepreneurs Under Uncertainty: An Economic Experiment in China . *Management Science*, *59*(7), 1671–1687.

Jovanovic, B. (1979). Job Search and the Theory of Turnover. *Journal of Political Economy*, *87*(5), 972–990.

Kahn, B. & Sarin, R. (1988). Modelling Ambiguity in Decisions Under Uncertainty. *Journal of Consumer Research*, *15*(2), 265–272.

Klein, D., Kotov, R., & Bufferd, S. (2011). Personality and Depression: Explanatory Models and Review of the Evidence. *Annual Review of Clinical Psychology*, *7*(1), 269–295.

Knight, F. (1921). *Risk, Uncertainty, and Profit*. New York: Houghton Mifflin.

Konon, A. (2016). Career Choice under Ambiguity. Paper presented at the Annual Conference of the German Economic Association (Verein für Socialpolitik) 2016.

Kull, S., Ramsay, C., & Lewis, E. (2003). Misperceptions, the Media, and the Iraq War. *Political Science Quarterly*, *118*(4), 569–598.

Lazear, E. (2005). Entrepreneurship. *Journal of Labor Economics*, *23*(4), 649–680.

Lazear, E. (2009). Firm-specific Human Capital: A Skill-weights Approach. *Journal of Political Economy*, *117*(5), 914–940.

Lopez, J. & Servén, L. (2006). A Normal Relationship? Poverty, Growth and Inequality. *World Bank Policy Research Working Paper*, *3814*.

Maccheroni, F., Marinacci, M., & Ruffino, D. (2013). Alpha as Ambiguity: Robust Mean-variance Portfolio Analysis. *Econometrica*, *81*(3), 1075–1113.

MacDonald, G. (1988). The Economics of Rising Stars. *American Economic Review*, *78*(1), 155–166.

Manso, G. (2016). Experimentation and the Returns to Entrepreneurship. *Review of Financial Studies*. forthcoming.

Mansour, H. (2012). Does Employer Learning Vary by Occupation? *Journal of Labor Economics*, *30*(2), 415–444.

Minnesota Population Center and State Health Access Data Assistance Center (2016). *Integrated Health Interview Series: Version 6.21*. Minneapolis: University of Minnesota. www.ihis.us.

Minniti, M. (2005). Entrepreneurship and Network Externalities. *Journal of Economic Behavior & Organization*, *57*(1), 1–27.

Newey, W. (1987). Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables. *Journal of Econometrics*, *36*(3), 231–250.

Paruchuri, S. & Ingram, P. (2012). Appetite for Destruction: The Impact of the September 11 Attacks on Business Founding. *Industrial and Corporate Change*, *21*(1), 127–149.

Paul, K. & Moser, K. (2009). Unemployment Impairs Mental Health. *Journal of Vocational Behavior*, *74*(3), 264–282.

Poschke, M. (2013). Who Becomes an Entrepreneur? Labor Market Prospects and Occupational Choice. *Journal of Economic Dynamics and Control*, *37*(3), 693–710.

Quatraro, F. & Vivarelli, M. (2015). Drivers of Entrepreneurship and Post-entry Performance of Newborn Firms in Developing Countries. *The World Bank Research Observer*, *30*, 277–305.

Rothschild, M. (1974). A Two-Armed Bandit Theory of Market Pricing. *Journal of Economic Theory*, *9*(2), 185–202.

Soroka, S. & McAdams, S. (2015). News, Politics, and Negativity. *Political Communication*, *32*(1), 1–22.

Soto, C., John, O., Gosling, S., & Potter, J. (2011). Age Differences in Personality Traits from 10 to 65. *Journal of Personality and Social Psychology*, *100*(2), 330–348.

Staiger, D. & Stock, J. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*(3), 557–586.

Stock, J. H. & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In D. Andrews & J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80–108). Cambridge: Cambridge University Press.

Strömberg, D. (2004). Radio's Impact on Public Spending. *Quarterly Journal of Economics*, *119*(1), 189–221.

Stuetzer, M., Goethner, M., & Cantner, U. (2012). Do Balanced Skills Help Nascent Entrepreneurs to Make Progress in the Venture Creation Process? *Economics Letters*, *117*(1), 186 – 188.

Terza, J., Basu, A., & Rathouz, P. (2008). Two-stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. *Journal of Health Economics*, *27*(3), 531–543.

The Economist (1997). Adventures with Capital. *The Economist*, (25 January 2016), p. 17.

Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies*, *25*(2), 65–86.

Verbruggen, M., van Emmerik, H., Van Gils, A., Meng, C., & de Grip, A. (2015). Does Early-career Underemployment Impact Future Career Success? A Path Dependency Perspective. *Journal of Vocational Behavior*, *90*(1), 101–110.

Vereshchagina, G. & Hopenhayn, H. (2009). Risk Taking by Entrepreneurs. *American Economic Review*, *99*(5), 1808–1830.

Wagner, J. (2006). Are Nascent Entrepreneurs 'Jacks-Of-All-Trades'? A Test Of Lazear's Theory of Entrepreneurship with German Data. *Applied Economics*, *38*(20), 2415–2419.

Wagner, J. (2007). What a Difference a Y Makes: Female and Male Nascent Entrepreneurs in Germany. *Small Business Economics*, *28*(1), 1–21.

World Bank Group (2015). www.doingbusiness.org/methodology.

Yatchew, A. & Griliches, Z. (1985). Specification Error in Probit Models. *Review of Economics and Statistics*, *67*(1), 134–139.

Zingales, L. (2000). In Search of New Foundations. *The Journal of Finance*, *55*(4), 1623–1653.

Appendix A

This appendix presents proofs of Lemma 1 and 2, and Proposition 1, 2, and 3.

A.1 Choice probabilities

Proof of Lemma 1. Note that $a_n = a_{\mathcal{S},n} + a_{\mathcal{W},n} = a_{\mathcal{S},0} + a_{\mathcal{W},0} + n$ such that a_n is deterministic. The binomial distribution can be constructed out of N i.i.d. draws from a Bernoulli distribution. Put differently, if

$$q = \sum_{n=1}^{N} q_n^*$$

where $q_1^*, \ldots, q_N^*$ are i.i.d. draws from Bernoulli(ϕ), then $q \sim$ Binomial(N, ϕ). Consider an arbitrary period. Assume that we selected and observed self-employment. The payoff is $\omega_{\mathcal{S}} \in \{0, 1\}$ and has distribution Bernoulli($\phi_{\mathcal{S}}$). The parameter $a_{\mathcal{S}}$ will be updated by adding $\omega_{\mathcal{S}}$, while the parameter $a_{\mathcal{W}}$ will be updated by adding $1 - \omega_{\mathcal{S}}$. Assume that we selected and observed wage work. The payoff is $\omega_{\mathcal{W}} \in \{0, 1\}$ with distribution Bernoulli($\phi_{\mathcal{W}}$). $a_{\mathcal{S}}$ is updated by adding $1 - \omega_{\mathcal{W}}$, whereas $a_{\mathcal{W}}$ is updated by adding $\omega_{\mathcal{W}}$. Let $\omega^* = 1 - \omega_{\mathcal{S}}$. It can be shown that $\omega^* \sim$ Bernoulli($\phi_{\mathcal{W}}$). The moment generating function of $\omega_{\mathcal{W}}$ is $M_{\omega_{\mathcal{W}}}(t) = \exp(t)\phi_{\mathcal{W}} + (1 - \phi_{\mathcal{W}})$. Consider the moment generating function of ω^*:

$$\begin{aligned} M_{\omega^*}(t) &= \mathbb{E}\left[\exp(t\omega^*)\right] = \exp(t)\mathbb{E}\left[\exp(-t\omega_{\mathcal{S}})\right] \\ &= \phi_{\mathcal{S}} + \exp(t)(1 - \phi_{\mathcal{S}}) = \exp(t)\phi_{\mathcal{W}} + (1 - \phi_{\mathcal{W}}) \end{aligned}$$

Hence, $\omega_{\mathcal{W}}$ can be replaced by $1 - \omega_{\mathcal{S}}$, as both have the same distribution. Furthermore:

$$\sum_{n=1}^{N} \omega_n^* = \sum_{n=1}^{N} (1 - \omega_{\mathcal{S},n}) \sim \text{Binomial}(N, \phi_{\mathcal{W}})$$

Now, let $q_{\mathcal{S},n} \sim$ Binomial($n, \phi_{\mathcal{S}}$) such that $a_{\mathcal{S},n} = a_{\mathcal{S},0} + q_{\mathcal{S},n}$ and $a_{\mathcal{W},n} = a_{\mathcal{W},0} + n - q_{\mathcal{S},n}$. Rewrite $\mathfrak{u}_n$ and $\mathfrak{b}_n$ as follows:

$$\mathfrak{u}_n(q_{\mathcal{S},n}) = \gamma_n^* + \delta_n^* q_{\mathcal{S},n} \tag{A.17}$$

$$\gamma_n^* \equiv -\frac{n + a_{\mathcal{W},0} - a_{\mathcal{S},0}}{a_n}, \quad \delta_n^* \equiv \frac{2}{a_n} > 0$$

$$\mathfrak{b}_n(q_{\mathcal{S},n}) = \gamma_n + \delta_n q_{\mathcal{S},n} \tag{A.18}$$

$$\gamma_n \equiv \frac{a_{\mathcal{S},0}\theta_{\mathcal{W}} - (n + a_{\mathcal{W},0})\theta_{\mathcal{S}} - (a_n^2 + a_n)n - (a_n^2 + a_n)a_{\mathcal{W},0} + a_{\mathcal{S},0}a_n^2 + a_{\mathcal{S},0}a_n}{a_n^2(a_n + 1)}$$

$$\delta_n \equiv \frac{\theta_{\mathcal{W}} + \theta_{\mathcal{S}} + 2a_n(a_n + 1)}{a_n^2(a_n + 1)} > 0$$

$\mathfrak{u}_n$ and $\mathfrak{b}_n$ are both strictly increasing in $q_{\mathcal{S},n}$, as $\mathfrak{u}_n'(q_{\mathcal{S},n}) = \delta_n^* > 0$ and $\mathfrak{b}_n'(q_{\mathcal{S},n}) = \delta_n > 0$, and invertible. We are interested in the probabilities to select wage work given by $\mathbb{P}(\mathfrak{u}_n \leqslant 0)$ and $\mathbb{P}(\mathfrak{b}_n \leqslant 0)$. Let $H(x; n, \phi)$ denote the cumulative distribution function of the binomial distribution given parameters n and ϕ. Using the properties $\mathfrak{u}_n$ and $\mathfrak{b}_n$, it is easy to establish that

$$\mathbb{P}(\mathfrak{u}_n \leqslant 0) = H(\tau_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}}), \quad \tau_n^{\mathfrak{u}} \equiv \mathfrak{u}_n^{-1}(0) = -\frac{\gamma_n^*}{\delta_n^*} \tag{A.19}$$

and

$$\mathbb{P}(\mathfrak{b}_n \leqslant 0) = H(\tau_n^{\mathfrak{b}}; n, \phi_{\mathcal{S}}), \quad \tau_n^{\mathfrak{b}} \equiv \mathfrak{b}_n^{-1}(0) = -\frac{\gamma_n}{\delta_n} \tag{A.20}$$

■

A.2 Optimality of unbiased utility

Proof of Lemma 2. We can derive the optimal strategy by slightly modifying the argument of Bradt et al. (1956). Let $\mathrm{s}^* = \{\mathfrak{d}_n^*\}_{n=1}^N$ where $\mathfrak{d}_n^* \in \mathbb{O}$ denotes the (unknown) optimal strategy. Furthermore, let

$$\mathcal{V}_N(a_{\mathcal{S},0}, a_{\mathcal{W},0}, \mathrm{s}) = \sum_{n=1}^{N} \mathbb{E}\left[\Omega(\mathfrak{d}_n)|a_{\mathcal{S},0}, a_{\mathcal{W},0}\right]$$

denote the expected payoff associated with some strategy s. Consider an arbitrary period $k > 0$. Let s_{N-k}^* denote the optimal strategy for the remaining $N-k$ periods and let $\mathcal{V}_{N-k}$ denote the corresponding expected payoff. Assume that we consider selecting self-employment in period $k+1$. The expected payoff in period $k+1$ is $\mu_{\mathcal{S},k}$. In case of a success, which occurs with probability $\mu_{\mathcal{S},k}$, $a_{\mathcal{S},k}$ is updated to $a_{\mathcal{S},k}+1$, while $a_{\mathcal{W},k}$ remains the same. In case of a failure in self-employment, which occurs with probability $1-\mu_{\mathcal{S},k}$, $a_{\mathcal{S},k}$ remains the same, whereas $a_{\mathcal{W},k}$ is updated to $a_{\mathcal{W},k}+1$. Hence, the expected payoff from selecting self-employment in period $k+1$ is

$$\begin{aligned}\mathcal{V}_{\mathcal{S}} \equiv\ & \mu_{\mathcal{S},k} + \mu_{\mathcal{S},k}\mathcal{V}_{N-k}(a_{\mathcal{S},k}+1, a_{\mathcal{W},k}, \mathrm{s}_{N-k}^*) \\ &+(1-\mu_{\mathcal{S},k})\mathcal{V}_{N-k}(a_{\mathcal{S},k}, a_{\mathcal{W},k}+1, \mathrm{s}_{N-k}^*)\end{aligned} \tag{A.21}$$

Given a similar line of reasoning, the expected payoff from selecting wage work in period $k+1$ is

$$\begin{aligned}\mathcal{V}_{\mathcal{W}} \equiv\ & \mu_{\mathcal{W},k} + \mu_{\mathcal{W},k}\mathcal{V}_{N-k}(a_{\mathcal{W},k}+1, a_{\mathcal{S},k}, \mathrm{s}_{N-k}^*) \\ &+(1-\mu_{\mathcal{W},k})\mathcal{V}_{N-k}(a_{\mathcal{W},k}, a_{\mathcal{S},k}+1, \mathrm{s}_{N-k}^*)\end{aligned} \tag{A.22}$$

We should select self-employment if $\mathcal{V}_{\mathcal{S}}$ is strictly larger than $\mathcal{V}_{\mathcal{W}}$; be indifferent if $\mathcal{V}_{\mathcal{S}}$ and $\mathcal{V}_{\mathcal{W}}$ are equal; and select wage work if $\mathcal{V}_{\mathcal{W}}$ is strictly larger than $\mathcal{V}_{\mathcal{S}}$. Note that this holds for an arbitrary period and is, therefore, a general prescription. Furthermore, notice that $1-\mu_{\mathcal{S},k} = \mu_{\mathcal{W},k}$ and $1-\mu_{\mathcal{W},k} = \mu_{\mathcal{S},k}$. Hence: We should strictly prefer self-employment if $\mu_{\mathcal{S},k} > \mu_{\mathcal{W},k}$; be indifferent if $\mu_{\mathcal{S},k} = \mu_{\mathcal{W},k}$; and strictly prefer wage work if $\mu_{\mathcal{W},k} > \mu_{\mathcal{S},k}$. This prescription is equivalent to the prescription made by relative unbiased utility. ■

A.3 Non-optimality of unbiased utility

Proof of Proposition 1. Note that the distribution function of the binomial is

$$H(\tau; n, \phi) = \sum_{k=0}^{\lfloor \tau \rfloor} \begin{pmatrix} n \\ k \end{pmatrix} \phi^k (1-\phi)^{n-k}$$

where $\lfloor \tau \rfloor$ is the greatest integer less than or equal to τ. Hence, a sufficiently large increase (decrease) in τ increases (decreases) H, by the properties of distribution functions. To assess a potential bias, use Lemma 1 to establish:

$$\tau_n^{\mathfrak{u}} - \tau_n^{\mathfrak{b}} = \begin{cases} > 0 & \text{if } \theta_{\mathcal{S}} < \theta_{\mathcal{W}} \\ = 0 & \text{if } \theta_{\mathcal{S}} = \theta_{\mathcal{W}} \\ < 0 & \text{if } \theta_{\mathcal{S}} > \theta_{\mathcal{W}} \end{cases} \quad \text{for all } n \tag{A.23}$$

According to Lemma 2, $\tau_n^{\mathfrak{u}}$ induces optimal behavior such that deviations from it constitute a bias reducing career successes. Case 1: If $\theta_{\mathcal{S}} < \theta_{\mathcal{W}}$, $\tau_n^{\mathfrak{u}} > \tau_n^{\mathfrak{b}}$ and biased utility has a potential bias against wage work, as $\mathbb{P}(\mathfrak{u}_n \leqslant 0) \geqslant \mathbb{P}(\mathfrak{b}_n \leqslant 0)$, respectively a potential bias for self-employment, as $1 - \mathbb{P}(\mathfrak{u}_n \leqslant 0) \leqslant 1 - \mathbb{P}(\mathfrak{b}_n \leqslant 0)$. Case 2: If $\theta_{\mathcal{S}} = \theta_{\mathcal{W}}$, there is no bias. Case 3: If $\theta_{\mathcal{S}} > \theta_{\mathcal{W}}$, $\tau_n^{\mathfrak{u}} < \tau_n^{\mathfrak{b}}$ and biased utility has a potential bias for wage work, as $\mathbb{P}(\mathfrak{u}_n \leqslant 0) \leqslant \mathbb{P}(\mathfrak{b}_n \leqslant 0)$, respectively a potential bias against self-employment, as $1 - \mathbb{P}(\mathfrak{u}_n \leqslant 0) \geqslant 1 - \mathbb{P}(\mathfrak{b}_n \leqslant 0)$. Note that

$$\frac{\partial}{\partial \theta_{\mathcal{S}}}(\tau_n^{\mathfrak{u}} - \tau_n^{\mathfrak{b}}) < 0, \qquad \frac{\partial}{\partial \theta_{\mathcal{W}}}(\tau_n^{\mathfrak{u}} - \tau_n^{\mathfrak{b}}) > 0$$

such that at some point (given a large enough $\theta_{\mathcal{S}}$ or $\theta_{\mathcal{W}}$), we have $\mathbb{P}(\mathfrak{u}_n \leqslant 0) > \mathbb{P}(\mathfrak{b}_n \leqslant 0)$ or $\mathbb{P}(\mathfrak{u}_n \leqslant 0) < \mathbb{P}(\mathfrak{b}_n \leqslant 0)$, i.e., the bias is relevant if either $\theta_{\mathcal{S}}$ or $\theta_{\mathcal{W}}$ is sufficiently large. ■

A.4 Media and behavior

Proof of Proposition 2. Denote $\tau_n^{\mathfrak{u}}$ affected by media $\mathfrak{m} > 0$ by $\mathring{\tau}_n^{\mathfrak{u}}$. The impact of media is given by

$$\frac{\partial}{\partial \mathfrak{m}} \mathring{\tau}_n^{\mathfrak{u}} = -\frac{1}{2} < 0 \tag{A.24}$$

Let $\mathring{\tau}_n^{\mathfrak{b}}$ denote $\tau_n^{\mathfrak{b}}$ given that self-employment is affected by $\mathfrak{m} > 0$. It follows that

$$\frac{\partial}{\partial \mathfrak{m}} \mathring{\tau}_n^{\mathfrak{b}} = B B_0^{-1} \tag{A.25}$$

$$\begin{aligned} B \equiv {} & -(\theta_{\mathcal{W}}^2 + (\theta_{\mathcal{S}} + \mathfrak{m}^2 + (2a_{\mathcal{W},0} + 2a_{\mathcal{S},0} + 2 - 2n)\mathfrak{m} - (2a_{\mathcal{W},0} + 2a_{\mathcal{S},0} + 1)n + a_{\mathcal{W},0}^2 + (2a_{\mathcal{S},0} + 2)a_{\mathcal{W},0} + a_{\mathcal{S},0}^2 \\ & + 2a_{\mathcal{S},0})\theta_{\mathcal{W}} + (3\mathfrak{m}^2 + (2n + 6a_{\mathcal{W},0} + 6a_{\mathcal{S},0} + 2)\mathfrak{m} + (2a_{\mathcal{W},0} + 2a_{\mathcal{S},0} + 1)n + 3a_{\mathcal{W},0}^2 + (6a_{\mathcal{S},0} + 2)a_{\mathcal{W},0} \\ & + 3a_{\mathcal{S},0}^2 + 2a_{\mathcal{S},0})\theta_{\mathcal{S}} + 2\mathfrak{m}^4 + (8a_{\mathcal{W},0} + 8a_{\mathcal{S},0} + 4)\mathfrak{m}^3 + (12a_{\mathcal{W},0}^2 + (24a_{\mathcal{S},0} + 12)a_{\mathcal{W},0} + 12a_{\mathcal{S},0}^2 \\ & + 12a_{\mathcal{S},0} + 2)\mathfrak{m}^2 + (8a_{\mathcal{W},0}^3 + (24a_{\mathcal{S},0} + 12)a_{\mathcal{W},0}^2 + (24a_{\mathcal{S},0}^2 + 24a_{\mathcal{S},0} + 4)a_{\mathcal{W},0} + 8a_{\mathcal{S},0}^3 + 12a_{\mathcal{S},0}^2 \\ & + 4a_{\mathcal{S},0})\mathfrak{m} + 2a_{\mathcal{W},0}^4 + (8a_{\mathcal{S},0} + 4)a_{\mathcal{W},0}^3 + (12a_{\mathcal{S},0}^2 + 12a_{\mathcal{S},0} + 2)a_{\mathcal{W},0}^2 + (8a_{\mathcal{S},0}^3 + 12a_{\mathcal{S},0}^2 \\ & + 4a_{\mathcal{S},0})a_{\mathcal{W},0} + 2a_{\mathcal{S},0}^4 + 4a_{\mathcal{S},0}^3 + 2a_{\mathcal{S},0}^2) \end{aligned}$$

where $B_0 > 0$ but the sign of B is ambiguous. However, it is easy to show that $B < 0$ if $\theta_{\mathcal{S}} > \mathring{\theta}_{\mathcal{S},n}$ where

$$\begin{aligned} \mathring{\theta}_{\mathcal{S},n} \equiv {} & (-\theta_{\mathcal{W}}^2 - ([-2\mathfrak{m} - 2a_{\mathcal{W},0} - 2a_{\mathcal{S},0} - 1]n + \mathfrak{m}^2 + 2(a_{\mathcal{W},0} + a_{\mathcal{S},0} + 1)\mathfrak{m} + a_{\mathcal{W},0}^2 + 2(a_{\mathcal{S},0} + 1)a_{\mathcal{W},0} \\ & + a_{\mathcal{S},0}^2 + 2a_{\mathcal{S},0})\theta_{\mathcal{W}} - 2\mathfrak{m}^4 - (8a_{\mathcal{W},0} + 8a_{\mathcal{S},0} + 4)\mathfrak{m}^3 - (12a_{\mathcal{W},0}^2 + (24a_{\mathcal{S},0} + 12)a_{\mathcal{W},0} + 12a_{\mathcal{S},0}^2 + 12a_{\mathcal{S},0} + 2)\mathfrak{m}^2 \\ & - (8a_{\mathcal{W},0}^3 + (24a_{\mathcal{S},0} + 12)a_{\mathcal{W},0}^2 + (24a_{\mathcal{S},0}^2 + 24a_{\mathcal{S},0} + 4)a_{\mathcal{W},0} + 8a_{\mathcal{S},0}^3 + 12a_{\mathcal{S},0}^2 + 4a_{\mathcal{S},0})\mathfrak{m} - 2a_{\mathcal{W},0}^4 \\ & - (8a_{\mathcal{S},0} + 4)a_{\mathcal{W},0}^3 - (12a_{\mathcal{S},0}^2 + 12a_{\mathcal{S},0} + 2)a_{\mathcal{W},0}^2 - (8a_{\mathcal{S},0}^3 + 12a_{\mathcal{S},0}^2 + 4a_{\mathcal{S},0})a_{\mathcal{W},0} - 2a_{\mathcal{S},0}^4 - 4a_{\mathcal{S},0}^3 - 2a_{\mathcal{S},0}^2) \end{aligned}$$

$$[\theta_{\mathcal{W}}+(2\mathfrak{m}+2a_{\mathcal{W},0}+2a_{\mathcal{S},0}+1)n+3\mathfrak{m}^2+(6a_{\mathcal{W},0}+6a_{\mathcal{S},0}+2)\mathfrak{m}+3a_{\mathcal{W},0}^2+(6a_{\mathcal{S},0}+2)a_{\mathcal{W},0}+3a_{\mathcal{S},0}^2+2a_{\mathcal{S},0}]^{-1}$$

It is straightforward to show that $\mathring{\theta}_{\mathcal{S},n}$ is strictly increasing in n. Moreover, it is easy to demonstrate that $\lim_{n\to\infty}\mathring{\theta}_{\mathcal{S},n} = \theta_{\mathcal{W}}$ such that

$$\mathring{\theta}_{\mathcal{S},1} \leqslant \mathring{\theta}_{\mathcal{S},n} < \theta_{\mathcal{W}}$$

Hence, if $\theta_{\mathcal{S}} \geqslant \theta_{\mathcal{W}}$, we have $B < 0$ and, consequently:

$$\frac{\partial}{\partial \mathfrak{m}}\mathring{\tau}_n^{\mathfrak{b}} < 0 \tag{A.26}$$

Now, given sufficiently intensive media, we must have

$$H(\mathring{\tau}_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}}) < H(\tau_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}}) \tag{A.27}$$

and

$$H(\mathring{\tau}_n^{\mathfrak{b}}; n, \phi_{\mathcal{S}}) < H(\tau_n^{\mathfrak{b}}; n, \phi_{\mathcal{S}}) \tag{A.28}$$

if $\theta_{\mathcal{S}} \geqslant \theta_{\mathcal{W}}$. Using Lemma 1, (A.27) and (A.28) imply

$$\mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{u}} = \mathcal{S}) > \mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{S}), \qquad \mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{u}} = \mathcal{W}) < \mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{W}) \tag{A.29}$$

$$\mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{b}} = \mathcal{S}) > \mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{S}), \qquad \mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{b}} = \mathcal{W}) < \mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{W}) \tag{A.30}$$

where $\mathring{\mathfrak{d}}_n$ is a choice affected by media and $\mathfrak{d}_n$ a choice without the influence of media. ■

A.5 Media and bias against self-employment

Proof of Proposition 3. Using Proposition 1 and given that $\theta_{\mathcal{S}}$ is sufficiently larger than $\theta_{\mathcal{W}}$, we have

$$H(\tau_n^{\mathfrak{b}}; n, \phi_{\mathcal{S}}) > H(\tau_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}})$$

As $\mathring{\tau}_n^{\mathfrak{b}}$ is strictly decreasing in $\mathfrak{m}$ if $\theta_{\mathcal{S}} > \theta_{\mathcal{W}}$ and $\lim_{\mathfrak{m}\to\infty}\mathring{\tau}_n^{\mathfrak{b}} = -\infty$, there exists only one $\mathfrak{m}_n^*$ solving

$$\mathring{\tau}_n^{\mathfrak{b}}(\mathfrak{m}_n^*) = \tau_n^{\mathfrak{u}}$$

Hence, for every n there exists an $\mathfrak{m}_n \in (0, \mathfrak{m}_n^*]$ such that

$$H(\tau_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}}) \leqslant H(\mathring{\tau}_n^{\mathfrak{b}}(\mathfrak{m}_n); n, \phi_{\mathcal{S}}) < H(\tau_n^{\mathfrak{b}}; n, \phi_{\mathcal{S}})$$

Put differently, for every n there always exists an $\mathfrak{m}_n \in (0, \mathfrak{m}_n^*]$ such that the bias against self-employment and for wage work is reduced:

$$\mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{S}) < \mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{b}}(\mathfrak{m}_n) = \mathcal{S}) \leqslant \mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{S}), \qquad \mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{W}) \leqslant \mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{b}}(\mathfrak{m}_n) = \mathcal{W}) < \mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{W})$$

Therefore, if $\mathfrak{m} \in (0, \mathfrak{m}^{**}]$ where $\mathfrak{m}^{**} \equiv \min\{\mathfrak{m}_1^*, \mathfrak{m}_2^*, \ldots, \mathfrak{m}_N^*\}$, there is at least one period n^* where $\mathbb{P}(\mathring{\mathfrak{d}}_{n^*}^{\mathfrak{b}} = \mathcal{S}) = \mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{S})$ and $\mathbb{P}(\mathring{\mathfrak{d}}_{n^*}^{\mathfrak{b}} = \mathcal{W}) = \mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{W})$, while for all the remaining periods $\mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{S}) \leqslant \mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{b}} = \mathcal{S}) \leqslant \mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{S})$ and $\mathbb{P}(\mathfrak{d}_n^{\mathfrak{u}} = \mathcal{W}) \leqslant \mathbb{P}(\mathring{\mathfrak{d}}_n^{\mathfrak{b}} = \mathcal{W}) \leqslant \mathbb{P}(\mathfrak{d}_n^{\mathfrak{b}} = \mathcal{W})$. Note, however, that media might also be too intensive such that $\mathfrak{m}_n > \mathfrak{m}_n^*$ resulting in $H(\mathring{\tau}_n^{\mathfrak{b}}(\mathfrak{m}_n); n, \phi_{\mathcal{S}}) < H(\tau_n^{\mathfrak{u}}; n, \phi_{\mathcal{S}})$, i.e., a bias against self-employment might be transformed into a bias *for* self-employment. ■

Appendix B

In this appendix, we describe our data.

B.1 Characteristics of micro panel

Table B.1 presents variable descriptions.

TABLE B.1. Variables in micro panel

Variable (source if not IHIS)	Description	Values
Consumption of articles about famous entrepreneurs (Google Trends)	Relative frequency of the search item "famous entrepreneurs" in the US regions	100 for state with highest frequency (SHF); all other values relative to SHF; numbers aggregated over time and normalized by 1,000
Number of disasters (Guha-Sapir et al. 2016)	Number of natural and complex disasters	Numerical
Self-employed	Individual is self-employed	1 = Self-employed; 0 = Not self-employed
Wage worker	Individual is a wage worker	1 = Worker; 0 = Not worker
Age	Individual's age	Numerical
Earnings	Total earnings during the previous calendar year	1 = $01 to $4999; 2 = $5000 to $9999; 3 = $10000 to $14999; 4 = $15000 to $19999; 5 = $20000 to $24999; 6 = $25000 to $34999; 7 = $35000 to $44999; 8 = $45000 to $54999; 9 = $55000 to $64999; 10 = $65000 to $74999; 11 = $75000 and over
Education	Educational attainment	1 = Never attended/kindergarten only; 2 = Grade 1; 3 = Grade 2; 4 = Grade 3; 5 = Grade 4; 6 = Grade 5; 7 = Grade 6; 8 = Grade 7; 9 = Grade 8; 10 = Grade 9; 11 = Grade 10; 12 = Grade 11; 13 = 12th grade, no diploma; 14 = High school graduate; 15 = GED or equivalent; 16 = Some college, no degree; 17 = AA degree: technical/vocational/occupational; 18 = AA degree: academic program; 19 = Bachelor's degree (BA, AB, BS, BBA); 20 = Master's degree (MA, MS, Med, MBA); 21 = Professional (MD, DDS, DVM, JD); 22 = Doctoral degree (PhD, EdD)
Years on job	Years on main or longest or last job	0 = Less than a year; $1, 2, 3, \ldots$ = Numerical value for number of years
Health	Health status	1 = Excellent; 2 = Very Good; 3 = Good; 4 = Fair; 5 = Poor

Effort	Felt everything an effort, past 30 days	0 = None of the time; 1 = A little of the time; 2 = Some of the time; 3 = Most of the time; 4 = All of the time
Feelings interfered with life	Feelings interfered with life, past 30 days	1 = A lot; 2 = Some; 3 = A little; 4 = Not at all
Hopeless	How often felt hopeless, past 30 days	0 = None of the time; 1 = A little of the time; 2 = Some of the time; 3 = Most of the time; 4 = All of the time
Nervous	How often felt nervous, past 30 days	0 = None of the time; 1 = A little of the time; 2 = Some of the time; 3 = Most of the time; 4 = All of the time
Restless	How often felt restless, past 30 days	0 = None of the time; 1 = A little of the time; 2 = Some of the time; 3 = Most of the time; 4 = All of the time
Sad	How often felt sad, past 30 days	0 = None of the time; 1 = A little of the time; 2 = Some of the time; 3 = Most of the time; 4 = All of the time
Worthless	How often felt worthless, past 30 days	0 = None of the time; 1 = A little of the time; 2 = Some of the time; 3 = Most of the time; 4 = All of the time
Born in US	Born in the United States	1 = Born in US; 0 = Not born in US
Female	Gender	1 = Female; 0 = Male
Non-white	Ethnicity	1 = Non-white; 0 = White
Full-time work	Usually work full time	1 = Usually full time; 0 = Usually not full time
Limitations	Has any activity limitation	1 = Limited in any way; 0 = Not limited in any way
Got dividends	Received income from dividends from stocks/funds, previous calendar year	1 = Yes; 0 = No
Got food stamps	Authorized to receive Food Stamps, last calendar year	1 = Yes; 0 = No
Got welfare	Received income from welfare/public assistance, previous calendar year	1 = Yes; 0 = No

In Table B.2, we show descriptive statistics. Correlations are provided in Figure B.1. Figure B.2 shows the shares of self-employed and wage workers conditional on regions and time.

B.2 Characteristics of macro panel

In Table B.3, we present descriptive statistics. In 48% of all countries, doing business is rela-

TABLE B.2. Descriptive statistics for micro panel
(a) Non-binary variables

Variable	Min	1st quartile	Median	Mean	3rd quartile	Max	SD
Consumption of articles about famous entrepreneurs	0.14	0.42	0.91	1.12	1.81	3.39	0.80
Number of disasters (countries and regions)	3.00	6.00	10.00	14.80	24.00	38.00	10.99
Age	18.00	25.00	36.00	37.48	48.00	64.00	13.34
Earnings	1.00	2.00	3.00	3.84	5.00	11.00	2.51
Education	1.00	14.00	16.00	15.65	18.00	22.00	3.06
Years on job	0.00	0.00	2.00	4.92	6.00	35.00	6.90
Health	1.00	2.00	2.00	2.40	3.00	5.00	1.04
Effort	0.00	0.00	1.00	1.36	2.00	4.00	1.27
Feelings interfered with life	1.00	2.00	3.00	2.96	4.00	4.00	0.98
Hopeless	0.00	0.00	0.00	0.63	1.00	4.00	0.99
Nervous	0.00	0.00	2.00	1.48	2.00	4.00	1.09
Restless	0.00	0.00	2.00	1.58	2.00	4.00	1.19
Sad	0.00	0.00	1.00	1.01	2.00	4.00	1.07
Worthless	0.00	0.00	0.00	0.48	1.00	4.00	0.9

(b) Binary variables

Variable	Share of individuals with characteristic
Self-employed	0.12
Wage worker	0.88
Born in US	0.85
Female	0.67
Non-white	0.24
Full-time work	0.26
Limitations	0.16
Got dividends	0.09
Got food stamps	0.07
Got welfare	0.02

tively easy. Figure B.3 shows the variation of country averages of the shares of self-employed and wage workers.

Table B.4 shows correlations. There is a strong negative correlation (–1.0) between the share of self-employed and the share of wage workers; i.e., it appears that most self-employed recruit themselves from the wage workers' group. Furthermore, less corruption is strongly positively correlated (0.8) with relative ease of doing business, i.e., doing business is easier in less corrupt societies—it might also be one reason for lower levels of corruption. Ease of doing business is negatively correlated with the share of self-employed (–0.6), but positively correlated with the share of wage workers (0.7).

Figure B.4 shows distributions of the dependent variables (original and transformed by the logit transformation) in our data set. Note that the shares of wage workers and self-employed

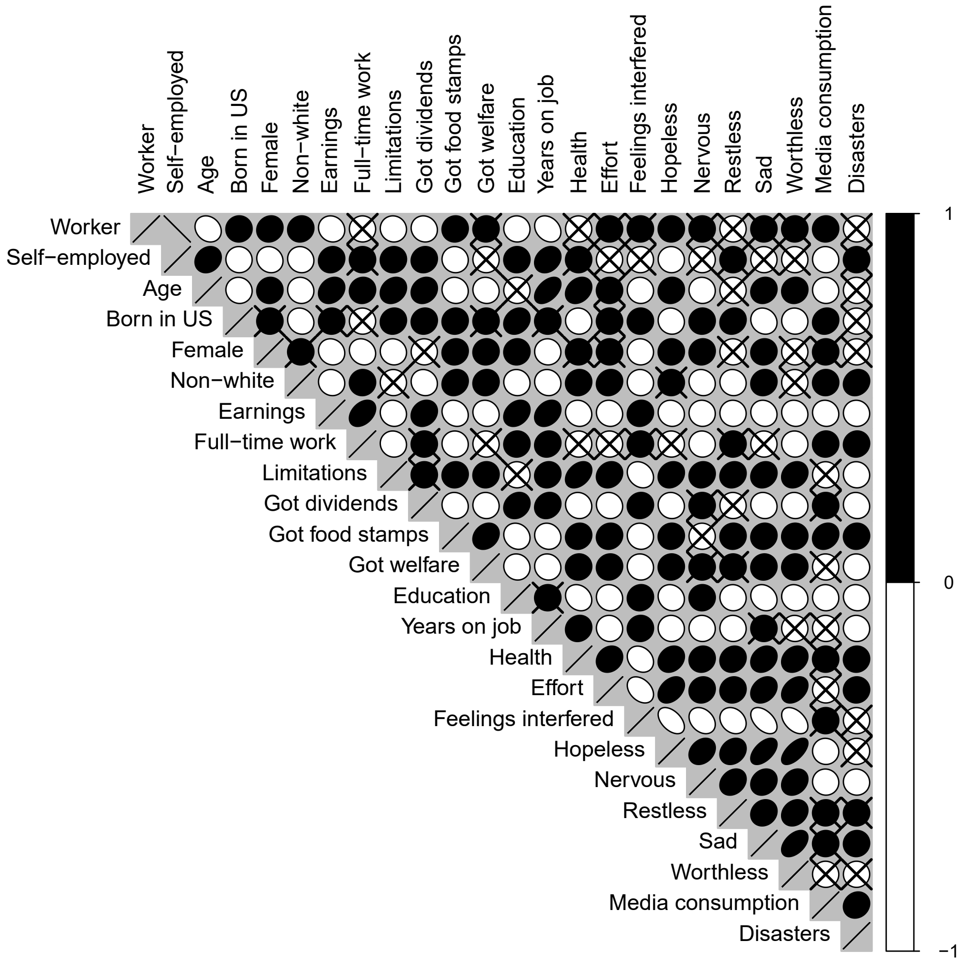


FIGURE B.1. Correlations in micro panel: Crossed out correlations are not significant at the 5%-level, while ellipses indicate strength (diagonal line is perfect correlation, whereas a perfect circle is no correlation) and direction (black is positive and white is negative correlation)

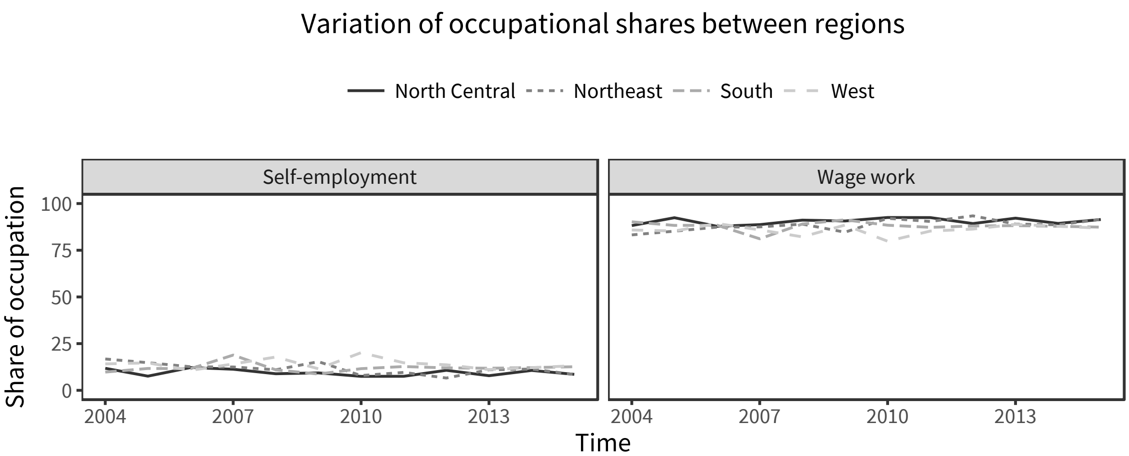


FIGURE B.2. Occupational shares

obey an asymmetric distribution and that skewness is not completely removed by the logit transformation (a reason to consider beta regression models).

Table B.3. Descriptive statistics for macro panel

Variable	Min	1st quartile	Median	Mean	3rd quartile	Max	SD
Media	19.00	49.00	57.00	57.64	67.00	88.00	14.64
Number of disasters (countries)	1.00	1.25	3.00	4.33	5.00	29.00	4.72
Share of self-employed	6.32	12.29	16.53	20.94	26.89	55.36	11.77
Share of wage workers	39.44	63.60	77.72	71.52	81.04	89.70	12.67
Fear of failure	0.15	0.28	0.32	0.33	0.36	0.58	0.07
Inflation	–6.01	1.88	3.21	4.53	5.52	34.93	5.34
GDP (per capita)	6187	15301	22695	25695	34551	77173	13152.52
GDP growth	–17.96	1.00	3.15	2.76	5.17	10.60	3.94
Real interest rate	–10.89	1.74	3.62	5.39	6.56	46.92	8.93
Lack of corruption	2.10	3.73	5.70	5.86	7.60	9.60	2.18

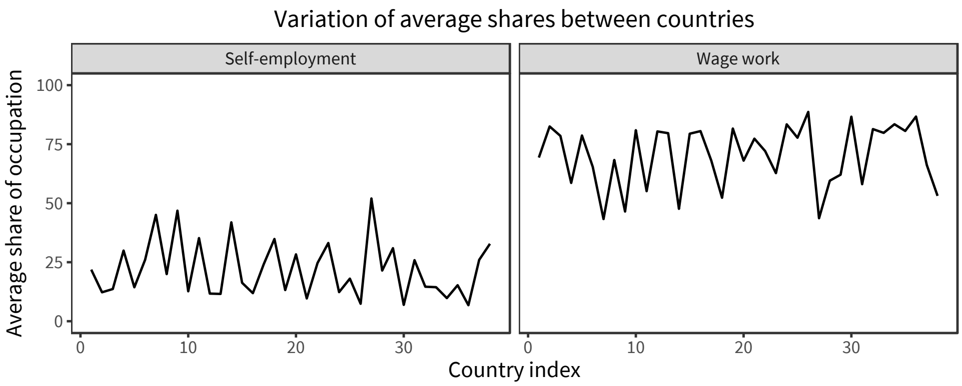


Figure B.3. Differences between countries in shares of self-employed and wage workers

Table B.4. Correlations in macro panel

	M	FOF	I	GDP	GDPG	RIR	COR	WWS	SES	DBE
FOF	–0.1									
I	0.1	0.1								
GDP	–0.1	–0.0	–0.4**							
GDPG	0.3**	–0.1	0.2**	–0.2**						
RIR	0.2**	0.1	–0.2**	–0.3**	–0.0					
COR	–0.0	–0.2**	–0.5**	0.8**	–0.1	–0.2**				
WWS	–0.3**	0.0	–0.3**	0.7**	–0.2**	–0.4**	0.7**			
SES	0.4**	–0.0	0.3**	–0.7**	0.3**	0.3**	–0.6**	–1.0**		
DBE	0.1	–0.0	–0.4**	0.7**	–0.1	–0.2**	0.8**	0.7**	–0.6**	
DIS	–0.5**	–0.1	–0.0	–0.1	–0.1	0.1	–0.0	0.1	–0.1	–0.2**

Notes: **Correlation is significant at the 5%-level; M = media reports intensity (GEM-based variable); FOF = fear of entrepreneurial failure; I = inflation; GDP = GDP per capita; GDPG = GDP growth; RIR = real interest rate; COR = lack of corruption; WWS = share of wage workers; SES = share of self-employed; DBE = doing business is relatively easy (dummy is 1 if yes); DIS = number of natural disasters in other country

FIGURE B.4. Distributions of dependent variables in macro panel

B.3 Stability of ease of doing business

In our empirical macro model, we assume that ease of doing business is sufficiently stable such that if doing business was relatively easy in 2010–2013, it was also relatively easy in 2003–2009. We, now, examine the stability of relative ease of doing business. We consider only countries where ease of doing business could be observed in all periods 2010–2013. Since the only variable of interest is ease of doing business, we do not have to ensure that all other variables are observed in the same period, meaning that data from a large number of countries (66 countries) is available. For each period $n = 2010, \dots, 2013$, we construct a group, denoted by $\mathcal{E}_n$, consisting of all countries where doing business was relatively easy in period n based on k-means clustering with two clusters. (In all countries not part of $\mathcal{E}_n$ doing business was relatively difficult.)

In Table B.5, we provide two measures of stability. First, the intersection with the previous

TABLE B.5. Stability of relative ease of doing business

Time	Number of countries where doing business is relatively easy	Intersection with previous period	Intersection with first period
2010	29		100%
2011	31	94%	94%
2012	32	97%	91%
2013	35	91%	83%

period is defined as

$$\text{Intersection with previous period}_n \equiv \frac{|\mathcal{E}_{n-1} \cap \mathcal{E}_n|}{|\mathcal{E}_n|} \in [0, 1]$$

and captures the number of countries where doing business was easy in period n and $n - 1$

relative to the number of countries where doing business was easy in period n. Second, the intersection with the first period is given by

$$\text{Intersection with first period}_n \equiv \frac{|\mathcal{E}_{2010} \cap \mathcal{E}_n|}{|\mathcal{E}_n|} \in [0, 1]$$

and captures the number of countries where doing business was easy in 2010 and period n relative to the number of countries where doing business was easy in period n.

If relative ease of doing business is stable, we expect both measures to be high. We find that the intersection with the previous period is never below 91%. The intersection with the first period is never below 83%. In particular, in 83% of all countries where doing business was relatively easy in 2013, doing business was also relatively easy in 2010. Consequently, relative ease of doing business is acceptably stable for our purposes.

Appendix C

Table C.1. Probit estimates of marginal effects in micro panel model without using IV, where dependent variable is choice dummy

	Self-employment		Wage work	
Variable	Coefficient	SE	Coefficient	SE
Consumption of articles about famous entrepreneurs	0.155**	(0.068)	–0.155**	(0.068)
Age	0.196***	(0.037)	–0.196***	(0.037)
Age^2	–0.002***	(0.000)	0.002***	(0.000)
Female†	–2.254***	(0.608)	2.254***	(0.608)
Born in US†	–0.182	(0.109)	0.182	(0.109)
Non-white†	–0.267	(0.176)	0.267	(0.176)
Full-time work†	–0.433***	(0.117)	0.433***	(0.117)
Earnings	0.004	(0.015)	–0.004	(0.015)
Got dividends†	0.323**	(0.135)	–0.323**	(0.135)
Got food stamps†	–0.145	(0.169)	0.145	(0.169)
Got welfare†	0.081	(0.278)	–0.081	(0.278)
Education	0.013	(0.019)	–0.013	(0.019)
Years on job	0.082***	(0.017)	–0.082***	(0.017)
Limitations†	0.050	(0.100)	–0.050	(0.100)
Health‡	0.024	(0.065)	–0.024	(0.065)
Effort	0.030	(0.031)	–0.030	(0.031)
Feelings interfered with life	–0.050	(0.045)	0.050	(0.045)
Hopeless	–0.084	(0.052)	0.084	(0.052)
Nervous	0.013	(0.033)	–0.013	(0.033)
Restless	0.079**	(0.034)	–0.079**	(0.034)
Sad	–0.030	(0.042)	0.030	(0.042)
Worthless	–0.074	(0.054)	0.074	(0.054)
North Central†	–0.340	(0.195)	0.340	(0.195)
Northeast†	–0.034	(0.200)	0.034	(0.200)
West†	0.024	(0.184)	–0.024	(0.184)
Constant	–5.341***	(0.707)	5.341***	(0.707)
Variable: Variance model		Coefficient	SE	
North Central†		–0.032	(0.073)	
Northeast†		–0.054	(0.083)	
West†		0.086	(0.072)	
Age		0.014***	(0.003)	
Female†		0.626***	(0.114)	
Non-white†		–0.087	(0.067)	
Health‡		–0.056**	(0.026)	
Education		0.000	(0.001)	
		10,851 obs.		

Notes: †Dummy variable; ‡increase indicates more health problems; ***significant at the 1%-level; **significant at the 5%-level; standard errors in parentheses are heteroskedasticity-consistent.

This appendix provides additional results and further information. Table C.1 presents results for a heteroskedastic probit without IV. In Table C.2, we present results for linear models without the use of an IV. Table C.3 shows the assignment yielding the strongest instrument in the macro

TABLE C.2. Estimates of marginal effects on transformed shares in macro panel model without IV, where dependent variable is transformed choice share

	Self-employment		Wage work	
Variable	Coefficient	SE	Coefficient	SE
Media attention for entrepreneurship	0.015***[**]	(0.003)	–0.011***[**]	(0.003)
Doing business is relatively easy†	–0.517***	(0.140)	0.359***	(0.102)
Fear of entrepreneurial failure	–0.294	(0.667)	0.362	(0.535)
Inflation	–0.016	(0.013)	0.012	(0.011)
GDP	0.000***	(0.000)	0.000***[**]	(0.000)
GDP growth	0.008	(0.012)	0.0.005	(0.010)
Real interest rate	0.003	(0.003)	–0.004	(0.003)
Lack of corruption‡	–0.047	(0.030)	0.074***	(0.021)
Constant	–1.278***[**]	(0.371)	0.428	(0.288)
		170 obs.		

Notes: †Dummy is 1 if yes and zero else; ‡increase indicates less corruption; ***significant at the 1%-level; **significant at the 5%-level; [***]significant at the 1%-level with country-level clustering; [**]significant at the 5%-level with country-level clustering; standard errors in parentheses are heteroskedasticity-consistent.

panel model. Table C.4 presents estimation results of linear probability models.

TABLE C.3. Assignment of countries resulting in strongest instrument

Country	Paired country with disasters	Country	Paired country with disasters
Argentina	Croatia	South Korea	Serbia
Australia	Latvia	Latvia	Hong Kong
Belgium	Greece	Malaysia	Jamaica
Brazil	Romania	Mexico	Venezuela
Canada	Trinidad and Tobago	Netherlands	Slovenia
Chile	Colombia	New Zealand	Pakistan
Colombia	Iran	Norway	Uruguay
Croatia	Guatemala	Peru	Italy
Ecuador	Germany	Poland	Australia
UK	Mexico	Romania	Spain
Greece	New Zealand	Russia	Turkey
Hong Kong	Sweden	Serbia	Peru
Hungary	USA	Singapore	Slovak Republic
Iran	Taiwan	Slovenia	Canada
Ireland	Norway	Sweden	Japan
Israel	Poland	Switzerland	France
Italy	Thailand	USA	Belgium
Jamaica	Switzerland	Uruguay	UK
Japan	Ireland	Venezuela	Hungary

A linear probability model corresponds to replacing $\mathfrak{L}(p_{j,n,i})$ in Equation (15) with $p_{j,n,i}$, such that the coefficient of media reflects the percentage-points effect on the probability of an

Table C.4. Marginal media effects according to IV linear probability models using macro panel, where dependent variable is original choice share

	Self-employment		Wage work	
Variable	Coefficient	SE	Coefficient	SE
Media attention for entrepreneurship§	0.005***[***]	(0.001)	–0.004***[***]	(0.001)
Doing business is relatively easy†	–0.091***	(0.025)	0.090***	(0.025)
Fear of entrepreneurial failure	–0.084	(0.098)	0.088	(0.100)
Inflation	–0.004	(0.002)	0.003	(0.003)
GDP	0.000**	(0.000)	0.000***[**]	(0.000)
GDP growth	–0.001	(0.002)	0.003	(0.002)
Real interest rate	0.000	(0.001)	0.000	(0.001)
Lack of corruption‡	–0.014***	(0.005)	0.018***[**]	(0.005)
Constant	0.161***	(0.059)	0.705***[***]	(0.059)
		170 obs.		

Notes: §Media attention is instrumented by number of natural disasters in other countries; †dummy is 1 if yes and zero else; ‡increase indicates less corruption; ***significant at the 1%-level; **significant at the 5%-level; [***]significant at the 1%-level with country-level clustering; [**]significant at the 5%-level with country-level clustering; standard errors in parentheses are heteroskedasticity-consistent.

occupation given that positive media attention for entrepreneurship increases by 1 percentage point.