Cross-Country Determinants of Early Stage Necessity and Opportunity-Motivated Entrepreneurship: Accounting for Model Uncertainty

Boris N. Nikolaev
Department of Entrepreneurship, Baylor University
Address: One Bear Place #98011, Waco, TX 76798
E-mail: boris_nikolaev@baylor.edu

Christopher J. Boudreaux
AR Sanchez Jr. School of Business, Texas A&M International University
Address: 5201 University Blvd., Laredo, TX 78041, USA
E-mail: cboudreaux@tamiu.edu

Leslie Palich
Department of Entrepreneurship, Baylor University
Address: One Bear Place #98011, Waco, TX 76798
E-mail: Les_Palich@baylor.edu

November 15, 2017

Abstract

In this cross-country study, we evaluate the robustness of 44 possible determinants of early stage opportunity-motivated entrepreneurship (OME) and necessity-motivated entrepreneurship (NME) that are broadly classified in five groups: (1) economic variables, (2) formal institutions, (3) cultural values, (4) legal origins, and (5) geography. The results, which are based on a representative world sample of as many as 73 countries, suggest that institutional variables associated with the principles of economic freedom are most robustly correlated with OME and NME. Our findings also identify net income inequality and Scandinavian legal origins as weakly robust predictors of both types of entrepreneurial activity. Furthermore, we find that log GDP per capita is only a weakly robust predictor of NME, but not OME. After accounting for model uncertainty, however, we do not find evidence to show that cultural values, geography, and legal origins are key indicators of OME and NME. We conclude by discussing the implications of our results for cross-country entrepreneurship research and practice.

Keywords: model uncertainty, entrepreneurship, model averaging, cross-country estimations
1. Introduction

Entrepreneurship has long been considered to be a crucial building block to the economic development (Kilby, 1971; North, 1990; Schumpeter, 1934), innovativeness (Terjesen, Hessels, & Li, 2016; Wong, Ho, & Autio, 2005), productivity (Bjørnskov & Foss, 2016), and growth (Audretsch, 2007; Audretsch, Keilbach, & Lehmann, 2006; Holcombe, 1998; Wennekers & Thurik, 1999) of nations. For this reason, governments around the world have been increasing the resources they commit to various programs and initiatives that are created to support entrepreneurial initiative and startup activity. It naturally follows that understanding the primary underpinnings and driving dynamics of this phenomenon is vitally important. But this begs the question: What if the research findings of the countless cross-country studies of the drivers of entrepreneurship have led to conclusions that might be misleading as several recent studies suggest (Bjørnskov & Foss, 2016; Su, Zhai, & Karlsson, 2016; Terjesen et al., 2016)? This would mean that the investment in these programs may not yield the expected results, and the opportunity costs would also be substantial since scarce funding would then have been diverted from more productive uses. As an attempt to address this concern, this study assesses the robustness of relationships between entrepreneurial action and its antecedent factors, as suggested by prior research, using a methodological approach that has not been applied to this body of research heretofore (Young & Holsteen, 2015).

The empirical literature on the cross-country determinants of entrepreneurship has grown considerably over the past decade. Previous studies have linked entrepreneurship to a number of important factors, including economic development (Audretsch & Acs, 1993; Audretsch et al.,

---

1 Terjesen, Hessels, & Li (2013) provide an excellent review of the comparative international entrepreneurship literature and conclude that out of 259 articles published in 21 leading entrepreneurship and management journals from 1989 to 2010, close to one-fourth of the articles are based on cross-country data.
2006; Baumol & Strom, 2007), cultural values (Tiessen, 1997; Welter, 2012), formal institutions (Bjørnskov & Foss, 2010; Kret & Sobel, 2005; McMullen, Bagby, & Palich, 2008; Nyström, 2008), and even post-materialistic values (Uhlaner & Thurik, 2010) and role models (Bosma, Hessels, Schutjens, Praag, & Verheul, 2012). Despite this breadth of research, however, cross-country empirical studies have so far provided less-than-complete insights into the mechanisms driving entrepreneurship (Bjørnskov & Foss, 2010; Su, Zhai, & Karlsson, 2016; Terjesen, Hessels, & Li, 2013). A major issue in this literature has been that of model uncertainty— that is, the use of a wide variety of explanatory variables has produced inconsistent empirical findings (Terjesen, Hessels, & Li, 2013). As Bjørnskov and Foss (2016, p. 301) note, “the entrepreneurship literature has not converged on a consensus on what to consider as a standard or even minimalist empirical specification.”

This is problematic because theory often does not provide enough guidance to select the proper empirical model (Heckman, 2005; Leamer, 1983; Raftery, 1995; Young & Holsteen, 2015). Rather, choosing a model is “difficult, fraught with ethical and methodological dilemmas, and not covered in any serious way in classical statistical texts” (Ho, Imai, King, & Stuart, 2007, p. 232). In fact, theory can be tested in many different ways, and since empirical findings depend on both

---

2 By model uncertainty, we mean the uncertainty that arises when analysts choose the true causal model that is then used in classical statistical tests. In the course of classical statistical analysis, authors often estimate a large number of models, but report only a handful of empirical estimations. These preferred models reflect only one “ad hoc” path in the modeling space (Leamer, 1985) and are likely to represent a non-random and unrepresentative sample of all plausible models (Sala-i-Martin et al., 2004). This is because significant findings that support the main hypotheses in a study are more likely to be published and therefore more likely to be reported. Young & Holsteen (2015), for example, review replication results reported in footnotes in two major sociological journals and find that out of 164 cases, not a single one of them reported results that failed to support the main findings. This is important because classical statistical theory assumes that only one (true) model is applied to a sample of data. If researchers, however, are uncertain which is the true model, this produces a much wider range of estimates than suggested by standard errors or confidence intervals associated with classical statistical methods. According to Pinello (1999) sampling uncertainty accounts for only 11% of the total variance in estimates while the remaining 89% is due to model specification (i.e., the choice of control variables in a model). We provide a more technical explanation of the problem of model uncertainty in the Data and Methodology section where we relate it to classical statistical tests, which account only for the uncertainty associated with the data sampling.
the data and the choice of model (Heckman, 2005), different models applied to the same set of data often produce different conclusions (Leamer, 1983). This can be clearly seen in the related cross-country literature on economic growth where more than 140 variables have been theoretically identified and found to be correlated with growth rates by different studies (Moral-Benito, 2010). However, in a set of classical robustness tests, Sala-i-Martin (1997) and Sala-i-Martin, Doppelhofer, and Miller (2004) found that the majority of these variables were consistently weak or non-significant, and many displayed significance in only 1 out of 1,000 regression models. Similar patterns have been documented in medicine (Ioannidis, 2005) and psychology (Simmons, Nelson, & Simonsohn, 2011) where the majority of published studies have been found to produce false positives instead of revealing robust correlations. Recent research has also pointed to a parallel credibility concern in strategic management journals as well (Bergh, Sharp, Aguinis, & Li, 2017). Such empirical heterogeneity, also known as model uncertainty, presents one of the biggest challenges for applied research (Durlauf et al., 2012; Glaeser 2008, Young 2009, Young and Holstene, 2015). It also emphasizes the great need for robustness analyses that can make empirical findings more compelling and less prone to non-robust, trivial, and false-positive results.

In this study, we identify robust cross-country determinants of entrepreneurship. To do this, we examine 44 variables as possible determinants of early-stage opportunity motivated entrepreneurship (OME) and necessity-motivated entrepreneurship (NME) for a representative sample of as many as 73 countries. The 44 explanatory variables can be broadly classified into five categories: (1) economic variables (e.g., log GDP per capita, net Gini coefficient, human capital), (2) formal institutions (e.g., economic freedom, democracy), (3) culture (e.g., individualism, social trust, religion), (4) legal systems, and (5) geography (e.g., latitude, tropics, or Scandinavian legal origins). The first four categories are common to many studies of cross-country phenomena—
e.g., see extensive reviews and other published research examining the determinants of life satisfaction (Bjørnskov, Dreher, & Fischer, 2008), international entrepreneurship (Terjesen et al., 2016), corruption (Serra, 2006) or economic growth (Sala-i-Martin et al., 2004). We added geographic predictors to the study because this factor provides information that is very different from the others we included and is common in economic growth research (Gallup, Sachs, & Mellinger, 1999).

At the outset, we would like to emphasize that the goal of our study is not to test any particular new theoretical notions, but to examine the robustness of a large number of variables identified by previous researchers. In that sense, we do not develop and test any novel hypotheses. As we argue later in the paper, one of the main reasons why the empirical literature has produced inconsistent findings and is running up against replicability challenges is precisely because of the publication bias that favors significant and novel results. Instead, here we put previous empirical findings under greater scrutiny by conducting a systematic “Global Sensitivity Analysis” based on an innovative methodology developed by Young and Holsteen (2015) and present a summary (i.e., the modeling distribution) of thousands of models that allows the reader to assess the weight of the evidence as opposed to a single (or a few) select findings. That is, we present results in the context of a distribution of plausible estimates and show the robustness of the estimated coefficients by taking into account not just the uncertainty inherent in the data, but also the uncertainty associated with the choice of control variables that are included in the model. While we are not aware of previous entrepreneurship studies using this methodology, this type of analysis has been successfully applied in many other disciplines (e.g., Bjørnskov et al., 2008; Sala-i-Martin, 1997; Serra, 2006) following the seminal work of Leamer (1983, 1985) on Extreme-Bound
Analysis. In this way, we answer recent calls in the entrepreneurship literature for robustness analysis (cf. Bjørnskov & Foss, 2016).

The remainder of this paper is organized as follows. Along with our review of literatures related to opportunity-motivated (OME) versus necessity-motivated entrepreneurial activity (NME) and the five variable categories above, we generate a series of research questions regarding the expected robustness of the relationships between the five factors and OME and NME. We then discuss the methods and analytical procedures used in our study, followed by a report of the results generated from the tests we conducted. Finally, we offer a discussion of the implications of our findings for cross-country research on entrepreneurship, as well as the limitations and practical applications of our investigation.

2. Drivers of Necessity- and Opportunity-Motivated Entrepreneurship

Entrepreneurship is a unique and challenging human endeavor that can manifest itself in a variety of different situations and contexts. Not surprisingly, previous studies have identified a long list of causes that motivate entrepreneurs to engage in the process of new venture creation (Carter, Gartner, Shaver, & Gatewood, 2003). Thus, defining and operationalizing entrepreneurship has been one of the greatest challenges in the field of entrepreneurship and is still hotly debated and largely unresolved (e.g., Parker, 2004; Wood, 2017). What makes studying entrepreneurship cross-nationally even more difficult is that reliable and consistent data for a representative number of countries is still generally lacking (Bjørnskov & Foss, 2016).

In this study, we use data from the Global Entrepreneurship Monitor (GEM), which defines entrepreneurship as “an attempt at a new business or new venture creation, such as self-employment, a new business organization, or the expansion of an existing business” (GEM, 2016).
This definition follows a long theoretical tradition recognizing that entrepreneurship refers not only to identifying a business opportunity, but also to acting on it through new business creation (Choi & Shepherd, 2004; Delmar & Shane, 2003; P. D. Reynolds, Carter, Gartner, & Greene, 2004; Ruef, Aldrich, & Carter, 2003).

More importantly, data from GEM allows us to differentiate between opportunity- and necessity-motivated entrepreneurship. Individuals engage in opportunity entrepreneurship when they perceive valuable business opportunities that can improve their lives. In that sense, the decision to engage in OME is not forced, but rather individuals are “pulled” into new business creation because of the promise of high extrinsic or intrinsic future rewards (e.g., Stoner & Fry, 1982; Gilad & Levine, 1986). Previous studies suggest that OME is associated with high-growth aspirations, export generation, and new product development (Hessels, Gelderen, & Thurik, 2008). Thus, the concept of OME is related to the Schumpeterian vision of entrepreneurship that ultimately leads to innovation and economic prosperity (McMullen et al., 2008). Individuals who engage in necessity-based entrepreneurship, on the other hand, do so because they have to, owing to the lack of other options. While necessity-motivated entrepreneurs range from street sellers to unemployed college graduates, what underlies their decision is that they are all “pushed” into entrepreneurship because they have no other employment prospects. By some estimates, over one billion people around the world today are considered necessity entrepreneurs (Brewer & Gibson, 2014). According to Margolis (2014), more than half of the workers in the developing world are self-employed, and a large proportion of these individuals engage in entrepreneurship because they must.

This distinction is important in our analysis for several reasons. First, a sizable literature at the micro-level shows that there are substantial differences between necessity and opportunity
entrepreneurs. For instance, one revealing study by Block & Wagner (2010) uses longitudinal data from Germany to show that the two groups differ significantly with respect to age, gender, and risk of becoming unemployed. They also find that opportunity entrepreneurs earn significantly higher incomes, which suggests that opportunity entrepreneurs have a stronger effect on economic growth. Adding to this profile, Reynolds, Camp, Bygrave, Autio, and Hay (2002) show that the two groups differ with respect to their growth aspirations, with 14 percent of opportunity entrepreneurs expecting to create more than 20 jobs in the future while only 2 percent of their necessity-motivated counterparts having similar growth aspirations. Furthermore, previous studies also show that necessity entrepreneurs are significantly more risk averse than opportunity-motivated entrepreneurs (Block, Sandner, & Spiegel, 2015).

As a second point of distinction, the rates of necessity and opportunity entrepreneurship at the country level are only weakly correlated (r=.27), suggesting that there are different causal mechanisms underlying each type entrepreneurship and that they should be considered separately (Reynolds et al., 2002). Third, because these two groups of entrepreneurs differ demographically, geographically, and occupationally (Reynolds et al., 2002), public policy regarding new ventures can differ substantially when it comes to each type of entrepreneurship (e.g., see Block and Wagner [2010] for the case of Germany). Finally, there are considerable differences in the levels of opportunity and necessity entrepreneurship across countries (Reynolds et al., 2002; GEM, 2016). Countries that are more developed economically tend to have higher ratio of opportunity-to-necessity entrepreneurs (e.g., the US with 1 to 6.4) compared to less developed countries (e.g., Nigeria at 2.5 to 1), though research indicates that even in developed countries the percentage of necessity-motivated entrepreneurs can be high. For example, according to GEM data from 2014,
the ratio of opportunity to necessity entrepreneurs in Germany is 2.32, which is similar to that in some developing countries such as Uganda, with a ratio of 2.9.

Taking all facets of this phenomenon into account, there are also concerns that looking at total rates of entrepreneurship or self-employment, as is common in the comparative entrepreneurship literature (Bjørnskov & Foss, 2010), can lead to misleading conclusions by erroneously implying that low levels of economic development, inefficient law enforcement, or uncertain regulatory environments are associated with more entrepreneurship (Henrekson & Sanandaji, 2014). For example, countries with high levels of total early-stage entrepreneurial activity (TEA), such as Burkina Faso and Ecuador (with TEA rates in excess of 30 percent), are often found at the bottom of international rankings on innovation (“Global Innovation Index 2016 Report,” 2016). To a great extent, much of the academic and policy debate so far has been about the differences, determinants, and merits of each of these two groups of entrepreneurs (e.g., see Block and Wagner, 2010, p.1). We therefore assess separately the linkages between prescribed determinants and the two forms of entrepreneurship (OME and NME).

2.1. Economic Factors

The empirical literature has identified a considerable list of variables as potential cross-country determinants of entrepreneurship. Traditionally, the focus has been on economic factors, such as economic growth, unemployment, human and physical capital accumulation, and technology. In one of the pioneering empirical studies, Audretsch and Acs (1993) found evidence that economic growth positively influences new start-up activity. Since their seminal study, a sizeable empirical literature has emerged that more directly examines the effects of macro-economic factors on entrepreneurial outcomes (Audretsch, 2007; Audretsch, Keilbach, & Lehmann, 2006; Wennekers
& Thurik, 1999; Freytag & Thurik, 2007), which often produces mixed empirical findings (Thai & Turkina, 2014).

An important insight from this literature is that factor endowments can play a crucial role in explaining entrepreneurial rates across countries. This logic is deeply rooted in neo-classical economics, which explains varying rates of economic growth and innovation across countries with differences in human and physical capital accumulation (Cass, 1965; Solow, 1956). Access to capital is a case in point. Because there is substantial variation in capital availability across countries (Djankov, McLiesh, & Shleifer, 2007), and because the lack of access to capital can critically hinder enterprise formation (De Soto, 2000), it naturally follows that this economic constraint is likely to be related to levels of productive entrepreneurship (Baumol, 2002). Margolis (2014), for instance, observes that restricted access to capital inhibits all forms of entrepreneurship, including necessity-based ventures and microenterprises, suggesting that physical capital would be an important variable in any cross-national study of new venture creation.

Similarly, human capital has also been increasingly linked to entrepreneurial success (for a review, see Unger, Rauch, Frese, & Rosenbusch, 2011). Previous studies, for example, suggest that human capital can facilitate the process of discovery of new entrepreneurial opportunities (Martin, McNally, & Kay, 2013) and later aid in successfully exploiting these opportunities during the start-up process (Bruns, Holland, Shepherd, & Wiklund, 2008). Human capital has been also connected to the accumulation of new knowledge that can lead to various advantages for firms (Bradley, McMullen, Artz, & Simiyu, 2012). Furthermore, human capital is closely associated with human adaptation (Weisbrod, 1962). This capacity is especially important in changing environments that require dynamic capabilities (Acemoglu & Autor, 2011), which certainly
characterize start-up situations in both necessity- and opportunity-driven businesses (Klein, Mahoney, McGahan, & Pitelis, 2013).

Despite these insights, in a recent review of the literature, Bjørnskov and Foss (2016) noted that factor endowments are rarely used as explanatory variables in cross-country studies of entrepreneurship. This creates risks for omitted variable bias and calls for careful robustness analysis. Bjørnskov and Foss (2016, p. 301) concluded that “the type of robustness studies pioneered in growth studies by Levine and Renelt (1992) and Sala-i-Martin (1997), in which researchers expose their main results and preferred specifications to extensive sets of additional, potentially influential factors and empirical alternatives, remain entirely absent in the literature on entrepreneurship.”

Recent work in labor economics and the economics of entrepreneurship conceptualizes self-employment as an occupational choice; that is, as an alternative to working for others (Klein, 2008; Parker, 2004). From this point of view, empirical evidence suggests that individuals are more likely to pursue entrepreneurial opportunities when the general economic conditions yield fewer higher-salary career alternatives (McMullen et al., 2008). In other words, entrepreneurial activity is likely to be greater when the opportunity costs of choosing self-employment are lower—i.e., when economic circumstances present fewer high-salary jobs. Thus, conceptually at least, both opportunity- and necessity-motivated entrepreneurship might be more prevalent in countries with lower levels of economic development and higher levels of unemployment (McMullen et al., 2008) where employment opportunities are more scarce.

Studies at the micro-level have also examined the link between opportunity and necessity-driven entrepreneurship and various economic factors, such as personal unemployment, education,
or earnings (e.g., see Block and Wagner, 2010). Some of the findings in this literature suggest that personal unemployment can lead to necessity entrepreneurship (e.g., Ritsilä & Tervo, 2002), and business start-ups borne out of necessity tend to be smaller, require less capital, and exhibit a slower pace of employment growth (Hinz & Jungbauer-Gans, 1999). Using data from West Germany, Pfeiffer & Reize (2000) find that such ventures do not necessarily perform worse in terms of employment growth and survival probability. Moreover, Block and Wagner (2010) show that opportunity entrepreneurs tend to earn significantly more income, which suggests that they likely have a greater impact on economic growth. These findings are noteworthy, but since our study is at the macro-level, we do not review the micro-level literature in greater detail.

Despite broad agreement that economic conditions play an important role in entrepreneurial action, models tested to date have produced inconsistent findings. As Terjesen et al. (2016, p.11) observe, results from pure economic indicators research have been less than consistent, partly “due to the diversity of measures of entrepreneurial and economic activity,” and partly due to the heterogeneity of samples and control variables employed in these analyses. Thai and Turkina (2014, p.2) make this point even more forcefully:

Although entrepreneurship scholars tend to agree on the categories of factors influencing entrepreneurship, their empirical studies have led to different conclusions with regard to the relative importance of each driver and at times to contrasting directions of influence. For example, several studies (e.g., Havrylyshyn, 2001; Kaufmann et al., 2006; Nystrom, 2008) show that good institutions and a high level of economic development and technology advancement are positively related to national rates of entrepreneurship. On the other hand, several other studies demonstrate that these same factors have a negative
relationship (e.g., Naude, 2009; Wong et al., 2005), a U-shaped relationship (Wennekers et al., 2005), or even no relationship at all (Van Stel et al., 2007).

This raises questions about the robustness of models assessing the relationship between economic factors and entrepreneurial activity. Though we maintain that the conventional wisdom about the relationship is justified, given the strength of supporting theory and the overall leaning of empirical evidence from previous research, this is far from assured. For this reason, we offer our first research question:

RQ1: *Is the relationship between general economic conditions and levels of entrepreneurial activity (OME and NME) robust across various models and tests of the relationship?*

### 2.2. Formal Institutions

While most neo-classical growth models treat the production process as a “black box” in which quantities of labor, capital, and technology are combined to produce new goods and services, more recent growth models emphasize the role of institutions, such as competitive markets, the banking system, or the structure of property rights as “fundamental” causes of entrepreneurship and ultimately economic growth and prosperity (North, 1990; Baumol, 1996; Acemoglu, Johnson, & Robinson, 2005; Campbell et al., 2012; Bennett & Nikolaev, 2016). This is because institutions, broadly defined as the “rules of the game in society” (North, 1991), reduce uncertainty in human interactions and structure the relative rewards from different productive and unproductive activities that can influence the allocation of entrepreneurial talent in the economy (Baumol, 1990).

Because entrepreneurship is a process that takes place over time (Mises, 1949), and because the future is unknowable, entrepreneurial action is inherently uncertain (McMullen & Shepherd, 2006). As “humanly devised constraints” on action (North, 1990, p. 3), formal institutions help to
ensure successful market transactions by reducing uncertainty in human interactions (Gwartney, Lawson, & Hall, 2016). For instance, when the legal rules are vague, contradictory, and constantly changing and law enforcement is weak, formal institutions can create more uncertainty instead of alleviating it. Under such conditions, it is difficult for entrepreneurs to forecast, plan, and engage in essential activities necessary to managing their ventures successfully. Similarly, when regulations are numerous and often changing, both public officials and entrepreneurs have a more difficult time navigating through the legal uncertainty (Boudreaux, Nikolaev, & Klein, 2017).

Institutions can also influence the distribution of entrepreneurial talent to different sectors of the formal and informal economies by shaping the relative rates of return from various productive activities, such as innovation, or unproductive ones, such as lobbying (Baumol, 1996; Sobel, 2008). In countries where people feel secure about their property rights, the rule of law prevails, scarce resources are allocated through the market system, private incentives are aligned with social ones, and the inflation rate is low and stable, more entrepreneurial talent will be allocated towards the discovery, evaluation, and exploitation of productive market opportunities (Baumol, 1996; Murphy, Shleifer, & Vishny, 1990). McMullen et al. (2008, p. 878) explain how these forces can determine, at least in part, the level of entrepreneurial action that takes place in a country:

Economic institutions … encourage the convergence of subjective models of the world by providing preexisting market constructs through which people understand the environment and solve the problems they confront (North, 1990, p. 20). In addition, they influence motivation as people contemplate the perceived costs of transacting, which consists of assessing the costs of measuring the valuable attributes of what is being exchanged and the costs of protecting rights and policing and enforcing agreements.
Thus, institutional factors such as a well-developed system of laws and property rights, streamlined and transparent business startup processes, and efficient economic regulation can promote opportunity-motivated entrepreneurial activity (Thai & Turkina, 2014). Numerous studies provide general support for many of these broadly held assertions (Bjørnskov & Foss, 2016; Grilo & Irigoyen, 2006; van Stel, Storey, & Thurik, 2007; Su et al., 2016).

Institutional conditions can be framed in many different ways, but in hundreds of studies they are conceptualized in terms of economic freedom. A recent review of the economic freedom literature (Hall & Lawson, 2014) revealed that an increasing share of these studies have explored potential connections to entrepreneurship. These investigations have found associations between entrepreneurial activity and a handful of institutional variables, including tax rates (Freytag & Thurik, 2007), labor regulations (Kreft & Sobel, 2005), and sound money practices (Bjørnskov & Foss, 2010). Empirical findings, however, have been inconsistent. On the one hand, lower tax rates (or relatively low freedom from government coercion) may hinder entrepreneurship (Freytag & Thurik, 2007). On the other hand, relatively high freedom from coercive labor restrictions and sound monetary policy may promote entrepreneurship (Bjørnskov & Foss, 2010).

McMullen et al. (2008) more specifically study the impact of formal economic institutions, measured by the Index of Economic Freedom published by the Heritage Foundation, on OME and NME. Contrary to Boudreaux et al. (2017), they hypothesized that underlying sub-components of the overall index—trade freedom, fiscal freedom, freedom from government, monetary freedom, investment freedom, labor freedom, property rights, business freedom, freedom from corruption, and financial freedom—will all be positively related to both OME and NME. They base these predictions on theory suggesting that, for example, high taxes dampen entrepreneurial interest as returns are reallocated to those who do not bear the risks of new ventures (Baumol, 2002),
government pricing interventions can lead to market distortions that make the costs and benefits of novel activities difficult to estimate (DiLorenzo, 2004), and government interference through wage and price controls discourage entrepreneurs by forcing them to accept elevated costs at a time when resources are tight and performance outcomes are difficult to project (Gwartney et al., 2016). Their research findings did not support their expectation in all cases and were significantly different than those of more recent research (Boudreaux et al., 2017).

Given the mixed results of studies on institutional conditions and the startup propensity of entrepreneurs across countries, questions about model robustness again come to mind. As was true with general economic conditions, there is a relatively rich body of theory and growing empirical evidence to indicate that the conceptualization has merit. But inconsistencies, some of which are highlighted above, suggest that the linkage needs to be assessed further. This leads to our second research question.

RQ2: Is the relationship between formal institutional conditions and levels of entrepreneurial activity (OME and NME) robust across various models and tests of the relationship?

2.3. Cultural Values

Autio, Pathak, & Wennberg (2013) point out that research from economics (Baumol, 1996; Greif, 1994), sociology (Aldrich, 2009), and international business (Stephan & Uhlaner, 2010) provides evidence to assert that national culture has some bearing on entrepreneurial activity across countries. This stands to reason, given the pervasive influence that cultures have on the development of informal societal values, norms, and traditions. This becomes clear from conceptualizations provided by researchers like Hayton, George, and Zahra (2002, p. 33):
Culture is defined as a set of shared values, beliefs, and expected behaviors (e.g., Herbig, 1994; Hofstede, 1980a). Deeply embedded, unconscious, and even irrational shared values shape political institutions as well as social and technical systems, all of which simultaneously reflect and reinforce values and beliefs. Cultural values indicate the degree to which a society considers entrepreneurial behaviors, such as risk taking and independent thinking, to be desirable. Cultures that value and reward such behavior promote a propensity to develop and introduce radical innovation, whereas cultures that reinforce conformity, groups interests, and control over the future are not likely to show risk-taking and entrepreneurial behavior (e.g., Herbig & Miller, 1992; Herbig, 1994; Hofstede, 1980a).

This notion of culture argues strongly for its inclusion in our understanding of cross-country entrepreneurship.

In a review of studies examining the role of culture as a driver of entrepreneurship across countries, Terjesen et al. (2013) found that Hofstede's (2001) perspective has had an influence that exceeds all others. In other words, the lion’s share of these studies (e.g., Del Junco & Brás-dos-Santos, 2009; Steensma, Marino, Weaver, & Dickson, 2000; Steensma, Marino, Weaver, and Dickson, 2000) have used Hofstede’s definition of culture and his five-dimension framework as a foundation for their research, perhaps because his contributions were derived from a research program that is still unparalleled in scope and is broadly shared and easy to understand. We should add, however, that other cultural frameworks have also been employed, including the GLOBE formulation (e.g., Stephan & Pathak, 2016), the World Values Survey (Pathak & Muralidharan, 2016) and original socio-cultural alternatives (e.g., Begley & Tan, 2001).
Though all of Hofstede’s dimensions have been included in cross-country entrepreneurship research, some have been deemed more apropos than others. For example, a number of researchers (e.g., Autio et al., 2013; Steensma et al., 2000a; Steensma et al., 2000b) chose to include only Individualism-Collectivism, Uncertainty Avoidance, and Masculinity-Femininity (or a close substitute) in their models and analyses. Autio and his colleagues (2013, p. 4) provided a justification for this, observing that these cultural indicators should be “particularly salient influences, because they resonate with the individualism, proactiveness, competitive orientation, innovativeness, and risk-taking commonly ascribed to entrepreneurial behaviors (Lumpkin & Dess, 1996).” Terjesen et al. (2013) observe that this research has indeed revealed some noteworthy findings. For example, individualism has been shown to promote startup effort (Baughn & Neupert, 2003; Mitchell, Smith, Seawright, & Morse, 2000) and entrepreneurship in the corporate setting (Morris, Davis, & Allen, 1994), and entrepreneurial collaboration seems to be more prevalent in societies that are more feminine and are uncertainty avoiding (Steensma et al., 2000a, 2000b).

These results are encouraging. However, hypothesized relationships have not always stood the test of empirical examination. Counter to the conclusions reported above, Wennekers, Thurik, Stel, and Noorderhaven (2007), found that individuals were actually more likely to engage in entrepreneurship, not less, when country-level uncertainty avoidance is high. More importantly, recent research in economics (Gorodnichenko & Roland, 2011b, 2011a) suggests that among all cultural dimensions, only individualism-collectivism has a significant effect on long-term growth and entrepreneurship. Research in political economy further points out that the link between cultural values, such as individualism, and economic growth and entrepreneurship works entirely through the channel of good governance (Kyriacou, 2016). That is, once we control for the quality
of formal institutions, the significant relationship between culture and economic growth completely disappears. These studies undermine many previous empirical findings and call for further robustness analysis. For these reasons, we propose our next research question:

RQ3: *Is the relationship between national culture and levels of entrepreneurial activity (OME and NME) robust across various models and tests of the relationship?*

### 2.4. Legal Origins

Though the three factors mentioned above—economic, institutional, and cultural conditions—have received significant attention in cross-country studies of entrepreneurship, we believe there is good reason to consider others. Specifically, we include here a discussion of legal origins and, in turn, geography because these have also been well studied and may indeed have an influential role in the shaping of startup intentions and growth outcomes for new or smaller ventures.

In the past couple of decades, social scientists have produced a considerable body of theoretical and empirical research suggesting that the historical origins of a country’s legal system and regulatory policies can affect a broad range of developmental outcomes (e.g., see La Porta, Lopez-de-Silanes, & Shleifer, 2008). In a series of papers, La Porta and coauthors, for instance, show that legal origins are strong predictors of financial institutions, government regulation, and judicial institutions. In turn, these factors are strong determinants of a variety of economic outcomes, such as the number of firms per capita, ownership concentration, the private credit-to-GDP ratio, employment rates, and contract enforcement.

The basic rationale for studying legal origins as determinants of entrepreneurship starts with the influential work of McNeill and McNeill (2003), who show how the transmission of information has shaped human societies over time. Information, they argue, can be transmitted
voluntarily via trade or migration flows or involuntarily through conquest and colonization. The information conveyed through these channels include technology, language, and religion, but also the law and legal systems (La Porta et al., 2008). While some types of information (such as technology) were often passed on voluntarily, legal traditions were typically transplanted involuntarily, and from a relatively small number of mother countries (Watson, 1974). Because of these tendencies, scholars assert that legal systems can be classified into major families, based on their historical background, characteristic modes of thought in legal matters, distinctive institutions, and ideology (Zweigert & Kötz, 1998). More specifically, these scholars identify two major legal traditions—common law and civil law—which are further broken down into sub-traditions—French, German, Socialist, and Scandinavian. It is not our aim here to discuss the origins, classification, and measurements of these different traditions since extensive reviews exist elsewhere (e.g., see La Porta et al., 2008).

More germane to our study, previous research suggests a strong correlation between legal origins and three broad categories of variables that are critical to entrepreneurship: (1) financial institutions and the development of capital markets, (2) government regulations and policies, and (3) judicial institutions. Some of the earliest studies on legal origins focused on the consequences of legal traditions on investor protection and financial development (La Porta, Lopez-de-Silanes, Shleifer, & Vishny, 1998). For example, in a series of papers, La Porta and co-authors show that legal origins are strong predictors of creditor rights, debt enforcement, prospectus disclosure, and shareholder protection from self-dealing by corporate insiders through corporate law (La Porta et al., 2008). They found that legal origins explained more than half of the variation in levels of debt enforcement and an anti-self-dealing index. In turn, these variables are strong predictors of the number of firms per capita, stock market-to-GDP ratio, ownership concentration, and private
credit-to-GDP ratio, all of which tend to be linked to various aspects of entrepreneurial activity at the country level (La Porta et al., 2008). Higher per capital income, innovation, and opportunity-motivated entrepreneurship, for instance, are typically associated with more developed financial markets as reflected by high stock market capitalization-to-GDP ratios, more firms per capita, and lower ownership concentrations.

Other studies focused on government regulation or public ownership of particular economic activities. For example, Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2002) examined the number of steps that entrepreneurs had to complete in order to start a business legally, which they found to vary from a total of two in Australia and Canada to 22 in the Dominican republic. Similarly, Botero, Djankov, Porta, Lopez-de-Silanes, and Shleifer (2004) studied the consequences of legal origins on labor regulations and the effects these directives have on labor force participation rates and unemployment. As La Porta et al. (2008) show, regulation entry and labor market regulations are significant determinants of corruption, employment in the informal economy, labor market participation rates, and unemployment, which can significantly affect opportunity costs across the formal and informal economies and influence the rates of OME and NME as previously discuss.

A third group of papers examined the effects of legal origins on the features of the judicial system and its effects on property rights and contract enforcement. As a case in point, Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2003) study the effect of legal origins on the efficiency of contract enforcement by courts (legal formalism), which turns out to be highly correlated with debt collection efficiency, contract enforcement, and property rights. These outcomes are all prerequisites for the efficient operation of markets and successful entrepreneurship. Overall, the weight of the evidence from these studies suggests that legal origins may play a central role in the
development of a variety of economic and political institutions over time, which can significantly affect rates of entrepreneurship.

Despite these observations, legal origins are rarely studied in the field of entrepreneurship, with most of the previous evidence coming from the fields of economics or finance. We conducted a search through the two leading entrepreneurship journals—the *Journal of Business Venturing* and *Entrepreneurship Theory and Practice*—to test this assertion, but this did not reveal a single article that directly examined the effects of legal origins on entrepreneurship.\(^3\) However, based on the abundance of studies that show a strong relationship between legal origins and a variety of market outcomes, such as the time to start a business, contract enforcement, employment rates, corruption, and ownership concentration, we maintain that these linkages are important and thus merit more careful examination. This leads to our next research question:

RQ4: *Is the relationship between legal origins and levels of entrepreneurial activity (OME and NME) robust across various models and tests of the relationship?*

### 2.5. Geographical Conditions

The central role of geography in economic development and innovation has long been recognized by social scientists (Marshall, 1890; Montesquieu, 1989; Smith, 1776). More recently, a substantial empirical literature in economics has suggested that geography plays an important role in shaping the distribution of income across countries and is a major driver of economic growth and prosperity (Gallup et al., 1999; Spolaore & Wacziarg, 2013). This thought goes back at least as far as Adam Smith (1776), who argued that “access to water reduced the cost of trade and gave merchants access to larger markets.” He noted that larger markets gave entrepreneurs

---

\(^3\) We searched for articles that contained the words “legal origins” or some variation of the term in their titles, abstract, or keywords.
incentive to specialize and innovate, which, in turn, stimulated the development of civilization along coastal areas where trade was easier. Even today, countries that are landlocked are, on average, much poorer than countries that have coastal access (Spolaore & Wacziarg, 2013). Of the 28 non-European landlocked economies in the world, there is not a single high-income country (Gallup et al., 1999). The poorest countries in South America are landlocked, including Bolivia and semi-landlocked Paraguay. Africa is the most landlocked continent in the world, with only one major river, the Nile, connecting countries within the continent. Not surprisingly, eleven of its fifteen landlocked nations are some of the poorest countries in the world, with per capita incomes of $600 or less (World Development Indicators, 2016). These are precisely the countries that are also found at the bottom of international rankings on innovation (Global Innovation Index 2016 Report, 2016) and have some of the highest rates of necessity-motivated entrepreneurship (GEM, 2016).

Poorer and less innovative countries also tend to be concentrated overwhelmingly in the tropics\(^4\) (Sachs, 2001). Economic underdevelopment in these regions can be partly explained by the negative effects of geography on two detrimental ecological handicaps: low agricultural productivity and the prevalence of infectious diseases. Tropical climates tend to be disadvantageous for photosynthesis, and the soil in these regions are prone to depletion due to heavy rainfall. Also, crops are often attacked by a host of pests and parasites that only thrive in hot climates (Masters & McMillan, 2001). Consequently, tropical plants tend to pack significantly fewer carbohydrates and are less nutritious. Perhaps more relevant for our study, a key determinant of the likelihood of increasing wealth over time has been the abundance of large domesticated animals, such as oxen or horses, which plays a key role in liberating significant portions of the

\(^4\) Tropical countries fall between 23.45 degrees North and South latitudes.
workforce from having to plow the land by hand. In tropical regions, however, domesticated animals historically have been the victims of a devastating array of diseases, such as trypanosomiasis, or sleeping disease (carried by the tsetse fly), which has been particularly harmful to domestic animals by making them lethargic or inactive (Swallow, 2000).

As with animals, humans in tropical regions have also been exposed to a terrifying array of diseases borne by insects and bacteria, such as malaria (Gallup et al., 1999; Sachs & Malaney, 2002). The prevalence of infectious diseases has greatly pushed up morbidity and mortality rates. In turn, unfavorable health and malnutrition conditions compound these effects by curbing the capacity of such societies to innovate and invest in human capital, which tends to impede technological development, diffusion of knowledge, and ultimately productivity (Gallup et al., 1999; Thornhill & Fincher, 2014). Virtually all of the low-income countries in the world today are simultaneously affected by at least five tropical diseases (Sachs, 2001). It follows logically that geographic latitude plays a major role in economic development, which explains its inclusion as a primary determinant in regression tests reported in economic growth studies (Spolaore & Wacziarg, 2013).

Despite the importance of geography to economic development and innovation, geographic controls are rarely used in cross-country entrepreneurship studies. Because of its potential role, as argued above, we think it is reasonable to consider the influence of geographic controls in our robustness analysis as well. This leads to our final research question:

RQ5: Is the relationship between geographic conditions and levels of entrepreneurial activity (OME and NME) robust across various models and tests of the relationship?

3. Data and Methodology

3.1. Model Uncertainty
The basic strategy of most cross-country entrepreneurship studies is to estimate a model of the following form:

\[ y_c = \alpha + \beta x_c + \sum_k \delta_k Controls_{k,c} + \epsilon_c \]  

Where \( y_c \) is some measure of entrepreneurial activity (e.g., self-employment, NME or OME, nascent entrepreneurship rates, etc.) and \( x_c \) is a variable of interest (e.g., the level of economic development). The researcher then picks \( k \) control variables, which are often assumed to increase the precision of the parameter estimates and decrease bias. The error term, \( \epsilon_c \), is often assumed to be uncorrelated with both \( x_c \) and \( y_c \). This is a critical assumption because omitted variables hidden in the error term can significantly bias the estimated coefficients (i.e., omitted variable bias). As a consequence, identifying which variables to include in the model is an essential strategy for causal identification in observational research (Heckman, 2005). But it is also one of the most common sources of uncertainty and ambiguity. In practice, the set of control variables rarely matches some well-established theoretical expectations, and adding or dropping control variables is a common practice in empirical research.

The problem is that when the model is not correctly specified, controlling for some variables but not others is as likely to reduce bias as it is to increase it (Clarke, 2005, 2009). This is because additional control variables leverage correlations with other omitted variables, which can intensify omitted-variable bias (Clarke, 2005). Controls can also leverage backward causal links with the outcome variable, producing reverse causation or selection bias (Elwert & Winship, 2014). In other words, when the “true” model is not known, control variables can have unpredictable consequences in practice. Thus point estimates often reported in papers represent a number of assumptions that usually capture only “one ad-hoc route through the thicket of possible models”
(Leamer, 1985, p. 208). When only one estimate is reported, these assumptions lead to “dogmatic priors” that the data must be analyzed with only one particular specification (Leamer, 2008).

Essentially, model uncertainty leads to the problem of asymmetric information between researchers and readers (Young, 2009). In the applied empirical research process, authors often “run many plausible models but in publication usually report a small set of curated model specifications” (Young & Holsteen, 2015, p. 4). Thus, researchers may know much more about the robustness of their results, but it is virtually impossible for readers to tell if the results are sensitive to model specification or if they are significant only somewhere in the model space (Ho et al., 2003). There are ways to address this potential pitfall. For example, researchers can employ multi-model analysis and put to scrutiny their results by estimating a large set of possible models as one way to relax model assumptions and demonstrate parameter robustness (Leamer, 1985; Raftery, 1995; Young & Holsteen, 2015).

Classical statistics is largely based on estimating uncertainty about the data—i.e., standard errors and confidence intervals only take into account uncertainty associated with random sampling. In other words, the estimated $b$ of the unknown parameter $\beta$ is based on chance since it is derived from a random sample. It is assumed that there are $K$ possible samples $\{S_1 \ldots S_k\}$ and each one of these random samples can generate a unique coefficient $\{b_1 \ldots b_k\}$. In repeated sampling, we would draw many samples and compute many estimates that would make up the sampling distribution. In this context, the sampling standard error, $\sigma_s = \sqrt{\frac{1}{K} \sum_{k=1}^{K} (b_k - b)^2}$, shows how much the estimate is expected to change if we draw a new sample. The critical assumption here, however, is that the true model is known.

If we admit that there is uncertainty about some aspects of the model—e.g., the set of control variables—then there will be more than $K$ estimates to consider, and the sampling distribution
alone will not convey the array of possible estimates. The modeling distribution, then, can be understood as analogous to the sampling distribution. With a \( j \) set of possible models \( \{M_1 \ldots M_j\} \) that might be applied to the data, there will be as many additional unique estimates of the unknown parameter \( \{b_1 \ldots b_j\} \). Here, \( \sigma_M \) is understood as modeling sampling error, which shows how much the coefficient is expected to change if we draw a random model from the list of plausible models. Thus, to fully account for uncertainty, we need to consider each possible sample \( \{S_1 \ldots S_k\} \), as well as each possible model \( \{M_1 \ldots M_j\} \). Taking the mean of these estimates, we can then calculate the mean estimate \( \bar{b} \) and total standard error, \( \sigma_T = \sqrt{\frac{1}{jk} \sum_{k=1}^{K} \sum_{j=1}^{J} (b_k - \bar{b})^2} \), which shows uncertainty about both the data and the model. Thus, rather than basing conclusions on sampling uncertainty alone, this approach incorporates model uncertainty as well (for an excellent discussion, see Young and Holsteen, 2015).

---

Insert Table 1 About Here
---

These issues can be clearly seen in our review of cross-country entrepreneurship studies above. We underscore this point in Table 1, which provides a summary of a number of cross-sectional studies that examine the effect of different variables on entrepreneurial activity. Though we present the list of studies in Table 1 for illustrative purposes only—not by any means as an extensive review of the literature—this sample of recent investigations reveals that this body of work lacks a conclusive underlying theoretical foundation, making it difficult to distinguish clearly which variables should be included in models to be tested. Some studies include only the level of economic development as a control variable, while others add a rich set of controls, including inequality or human capital. Some focus on cultural values while others explore formal institutions.
In short, there is a remarkable heterogeneity about the choice of controls that go into estimations of entrepreneurship outcomes across countries. In fact, three recent independent reviews of the comparative entrepreneurship literature (Bjørnskov & Foss, 2010; Su et al., 2016; Terjesen et al., 2016) have concluded that empirical findings to date have been mixed and confusing, maintaining that robustness tests are necessary to establish a better understanding of the underlying factors that drive cross-country differences in entrepreneurship.

A key step in testing model robustness is defining the model space—“the set of plausible models that analysts are willing to consider” (Young & Holsteen, 2015). In this paper, we focus on a model space defined by a set of control variables (Leamer, 1985; Raftery, 1995; Sala-i-Martin, 1997). Specifically, we examine a large number of plausible determinants of opportunity- and necessity-motivated entrepreneurship that are broadly classified according to the five general factors previously identified: (1) economic factors, (2) formal institutions, (3) cultural values, (4) legal origins, and (5) geographical conditions.

The downside of this approach is that it is computationally intensive. For instance, with 20 possible controls, there are more than a million unique possible models. Therefore, we pick 14 core variables from each of the five general factors above—log GDP per capita, income inequality (net), human capital, physical capital stock, employment, exchange rate, economic freedom, fiscal freedom, democracy, individualism, ethnolinguistic fractionalization, religion (Catholic, Protestant, Muslim, others), legal origins (Socialist, Great Britain, Scandinavian, French, and German), and geographic (latitude, tropics, and continent) controls. Because the procedure is computationally demanding, we treat religion, legal origins, and geographic controls as vectors, entering all variables in each vector together into individual regressions. These variables define our model space of 8,129 possible models. We then estimate each model and report our results as
a *distribution* of plausible estimates that show the robustness of the coefficient by taking into account not just the uncertainty inherent in the data, but also the uncertainty associated with the choice of control variables that are included in the model. In addition, we test several alternative variables.

### 3.2. Data

Because we use 44 variables from 14 different sources, Table 2 reports summary statistics and provides descriptions and sources of all variables used in this study. Data on entrepreneurship were collected from the GEM 2001-2015 Adult Population Survey (APS) Global Key Indicators file, which is freely available at http://www.gemconsortium.org/data. One of the key lessons from cross-country research on entrepreneurship has been the distinction between opportunity-motivated entrepreneurship (OME) and necessity-motivated entrepreneurship (NME). While numerous entrepreneurs around the world report being opportunity-driven—i.e., seeking to improve their situation either through increased independence or higher income—many others pursue entrepreneurship out of necessity. Using an overall measure of entrepreneurship, then, in a large representative sample of countries may produce confusing findings as we discussed in the theory section. Thus, our main variables of interest differentiate between the percentage of total early-stage OME and NME. We also focus on early-stage entrepreneurship because numerous studies find that small businesses in their early stages are largely considered to be one of the main drivers of economic growth, knowledge diffusion, and new job creation (Audretsch et al., 2006; van Praag & Versloot, 2007).

-------------------------------
Insert Table 2 About Here
-------------------------------
Data on entrepreneurship has only recently become available for a more representative set of countries. Most previous cross-country studies are based on small unrepresentative samples, primarily from developed Western economies. Since GEM data for each country are not available for every year, we use country averages over the timespan of the sample to maximize country coverage. This significantly increases our sample to 73 countries for most of our estimations. Since institutional, cultural, legal, and geographic variables are unlikely to change significantly over a short period of time (Alesina & Giuliano, 2015; Bjørnskov & Foss, 2016), we believe that this approach is appropriate. The correlation between the index of economic freedom in 2014 and the average of the index over 2001-2014 is close to perfect (r=0.94). One way to interpret our results, then, is as long-run trends as opposed to short-term fluctuations.

4. Comparison with Alternative Model Averaging Approaches and Meta-Analysis

Our study approach is clearly distinct from the alternatives that other researchers have employed. Traditional model-averaging approaches, for example, usually weigh the estimates either by model fit or by Bayesian priors. We follow instead recent advancements in the literature on model averaging (Young & Holsteen, 2015), which focus on the raw (unweighted) distribution of the estimated coefficients. In that sense, the focus is not on the “best” estimate, but on what estimates can be obtained from the data. The advantage of this approach is that it does not require making additional assumptions about the variables included in the model or their relative importance. For instance, obtaining high model fit in regression analysis is often associated with the inclusion of endogenous variables that are jointly determined with the outcome variables. But this requires the additional assumption that all control variables in the model are exogenous (Pearl, 2010; Sala-i-Martin, 1997). Similarly, weighing outcomes on the basis of Bayesian priors represents the researcher’s beliefs about model validity and favors a particular set of model
assumptions. As Young & Holsteen (2015) note, the goal of robustness analysis is to show how estimated results may change “under different beliefs about the correct model.” Since weighted estimates do not address the asymmetry of information between researchers and readers, we focus on the modeling distribution of unweighted estimates because they require the fewest assumptions about model specification.

Another common procedure to establish statistical significance (more accurately, to estimate the magnitude of effects) in studies that have conflicting results is meta-analysis. This approach allows researchers to systematically combine estimations from previous studies of similar type (e.g., randomized controlled trials) to reach a single conclusion that is statistically more robust than that of a single study. Conducting a meta-analysis, however, requires that the analysts make a number of additional assumptions. For example, identifying the appropriate set of studies to be included in the analysis tends to be a difficult and time-consuming process, which is based on additional rules adopted by the authors that are highly subjective (Stegenga, 2011). In this respect, a common problem with meta-analyses is the so called agenda-driven bias, which occurs when authors lean toward an economic, social, or political agenda and thus cherry-pick studies to be included in their analyses. Consistent with this, Roseman et al. (2011) found that information reflecting conflict of interest in underlying studies used in meta-analyses is rarely disclosed.

An important drawback in the context of our study is that meta-analyses rely on the critical assumption that both published and unpublished results are recognized in order to avoid publication bias that arises due to the so called “file drawer problem” (Rosenthal, 1979). This problem occurs because studies which show unsurprising, negative, or insignificant results are less likely to be published. In reality, uncovering and then incorporating information on unpublished studies, which can significantly change meta-analytic conclusions, is extremely challenging. Thus,
relying on the available body of published research, which is based disproportionately on a biased
collection of studies reporting significant findings, can lead to a serious base-rate fallacy. Indeed,
previous studies suggest that more than 25 percent of meta-analyses suffer from publication bias
(Ferguson & Brannick, 2012). Finally, an important drawback of meta-analytical studies has to do
with decisions about combining heterogeneous populations, which is also shaped by a number of
additional assumptions that authors make.

In contrast, the goal of the current analysis is to present the modeling distribution across a large
number of conceivable models, with the fewest possible assumptions about the model space from
a single sample. In that sense, readers are left to evaluate the weight of the evidence with respect
to a distribution of plausible estimates by evaluating several robustness statistics, such as
significance rate or sign stability, instead of being presented with the “best” possible estimate.
Finally, entrepreneurship scholars can easily incorporate the methodology we follow in this study
without the burden of collecting, synthesizing, and analyzing a large number of studies, which
involves a number of additional assumptions that increase the likelihood of researcher bias. While
not perfect, this provides a quick and easy way to present results in a more transparent way that
reduces the asymmetric information problem that exists between readers and analysts (Young,
2009).

5. Results

In this section, we present our sensitivity analysis. We follow the strategy outlined by Young
and Holsteen (2015) and report a number of statistics associated with the distribution of the
estimated coefficients from all 8,129 possible models in our model space, based on the 14 core
variables described in the previous section. In other words, we are asking whether the findings
associated with the estimated coefficients hinge on the sets of possible control variables, or do they hold regardless of the assumptions made about the model space (i.e., regardless of what control variables we include in the model).

Combining the mean estimate of \( \bar{b} \) and the total standard error, \( \sigma_T \), we report the “robustness ratio”: \( \frac{\bar{b}}{\sigma_T} \). This ratio is constructed as being analogous to a T-statistic and, following Young and Holsteens’s (2015) rule of thumb, we interpret values greater than two to reflect strong robustness. In addition, we report core summary statistics such as sign stability (the share of estimates that have the same sign) and significance rate (the share of models that report statistically significant coefficients). Raftery (1995) suggests, as an additional rule of thumb, that a significance rate of 50 percent sets a lower bound for “weak” robustness while a rate of 95 percent or higher indicates “strong” robustness. Thus, in all tables we rank variables according to their robustness as indicated by (1) the robustness ratio, (2) sign stbaility, and (3) significance rate. We consider an estimated coefficient to be “strongly robust” if it has (1) a robustness ratio \( > 2 \) or (2) a significance rate \( > 90 \) percent. We use the following rule for a “weakly robust” coefficient: (1) robustness ratio \( > 1.8 \) or (2) a significance rate \( > 50 \) percent. In both cases, in order to call a variable robust, we expect to see a sign stability of at least 90 percent. All other estimates are considered “non-robust.”

We report the robustness analysis in Table 3, which summarizes the results with respect to our economic controls, including log GDP per capita, factor endowments (human capital and physical capital stock), gross and net income inequality, and measures of employment. The findings in this table suggest that people who live in countries with higher levels of income inequality are less likely to engage in opportunity-motivated entrepreneurship and are more likely to engage in
entrepreneurial action out of necessity. The variable on income inequality, however, is only weakly robust, with a significance rate of 70 percent and a robustness ratio that implies significance at the 0.001 percent alpha level only in the case of NME. Importantly, we find that the level of economic development, measured by the log GDP per capita, which is one of the most frequently included variables in cross-country entrepreneurship regressions, displays (weak) robustness only with NME, but not with OME. Additionally, factor endowments, such as physical and human capital, as well as measures of employment and the proportion of people in the labor force who work in the industry or service sectors do not appear to be robustly correlated with either type of entrepreneurship.

These results, of course, as well as the results that follow, should be interpreted with caution. Our sensitivity analysis only suggests robust correlations and has nothing to say about the direction of causality. It is possible, for instance, that entrepreneurship drives inequality and not the other way around. In fact, the equality-efficiency trade-off is well-established in the economics literature (see Okun, 2015). It is generally accepted in economics that inequality plays an important efficiency role in the economy. Higher rewards generate productivity, but the newly created wealth is inevitably redistributed unequally in society as entrepreneurs enter new markets that generate tremendous wealth for themselves. Importantly, the variable on income inequality has not received much attention in the entrepreneurship literature, and our study identifies a fruitful avenue for future research that may provide a better understanding of the underlying mechanisms that drive cross-country differences in entrepreneurship.

--------------------------------------
Insert Table 4 About Here
--------------------------------------
Next, Table 4 shows the results with respect to formal institutions. In particular, we focus on one particular aspect of the institutional environment—economic freedom, a summary measure that is most commonly associated with institutional features such as limited government, freedom from corruption, open markets, sound monetary policy, freedom to trade, and lower regulatory burden. We choose to focus on economic freedom for two main reasons. First, most of the institutions-entrepreneurship literature so far has used the market logic inherent in the concept of economic freedom as a theoretical foundation (Su et al., 2014). Free market economies, for instance, reward effort and creativity with high social status, which encourages entrepreneurs to enter new markets generating extraordinary wealth through experimentation and innovation. There is also a growing empirical literature that has linked economic freedom to different entrepreneurial outcomes. Bjørnskov and Foss (2016, p. 298) provide an excellent overview of this literature and conclude that, despite promise, much of this literature is still in its infancy, and previous studies have “somewhat arrived at opposite conclusions.” In addition to economic freedom, we also test how the political institutions of democracy are correlated with economic freedom.

The main result reported in Table 4 is that economic freedom is found to be the most robust determinant of entrepreneurial activity across countries. Specifically, institutions associated with sound monetary policy, lower levels of corruption, and business and investment freedoms have the strongest and most robust correlation with entrepreneurial action among all variables tested in our study. Interestingly, while we find that higher levels of economic freedom are robustly and positively correlated with entrepreneurial activity, fiscal freedom is negatively and weakly correlated with NME. We do not find a robust relationship between democracy and entrepreneurship.

--------------------------------------

Insert Table 5 About Here
Table 5 presents the results from the robustness analysis of cultural values. A large body of research reports significant associations between cultural values and entrepreneurial activity (for a summary, see Terjesen, Hessels, & Li, 2016). A consistent finding in this literature is that individualistic values, most often measured by Hofstede’s cultural dimensions, are positively correlated with different measures of entrepreneurship (Baughn & Neupert, 2003; Gorodnichenko & Roland, 2011a, 2011b; Mitchell, Smith, Seawright, & Morse, 2000). Previous studies have also linked cultural values such as uncertainty avoidance and femininity to entrepreneurial outcomes (Steensma et al., 2000; Wennekers et al., 2007). None of Hofstede’s cultural dimensions, however, is robustly correlated with OME or NME in our analysis. If anything, the most robust variables in this category are the cultural dimensions of indulgence, masculinity, and social trust, which in the best case are only statistically significant in about a third of our estimated models. We find individualism at the bottom of the list, with a very low robustness ratio and significance rate of only 2 percent. We furthermore find that religion, measured by the proportion of people who belong to different major religions, is not a significant predictor of entrepreneurship.

Finally, in Table 6 we examine the effect of several geographic controls, such as the share of the population that lives in tropical areas, latitude, and continental dummies. For presentation purposes, we also include the results from legal origins in the same table (La Porta et al., 2008; La Porta, Lopez-de-Silanes, Shleifer, & Vishny, 2000). Geographic controls and legal origins are sporadically used in the entrepreneurship literature, but a long tradition in economics suggests that
geography and legal origins are important determinants of economic growth (Acemoglu et al., 2005; Bennett & Nikolaev, 2016a; La Porta et al., 2008). The only robust variable in this category is Scandinavian legal origins, which to our knowledge has not received much attention in the entrepreneurship literature either.

6. Discussion

The literature on the determinants of entrepreneurship has grown significantly over the past decade, identifying a large number of possible variables that can explain sizeable cross-country differences in entrepreneurial activity. However, as we discussed early on in the paper, empirical findings have been far from consistent, creating doubts about the robustness of previously established relationships (Bjørnskov & Foss, 2016; Su et al., 2016; Terjesen et al., 2016; Thai & Turkina, 2014). One possible explanation for this empirical heterogeneity is model uncertainty, and several independent reviews of the literature recognize that such heterogeneity has led to “limited insights” into the mechanisms that drive entrepreneurship (Bjørnskov & Foss, 2016; Su et al., 2016; Terjesen et al., 2016). Despite the tacit acknowledgement of the problem of model uncertainty, which is manifested by additional regression results with alternative specifications that scholars often report, traditional econometrics to a great extent has failed to exploit this systematic problem with applied empirical research. An exception here is the limited work on model uncertainty and model averaging pioneered by Leamer (1983, 1985), Raftery (1995), Sala-i-Martin (1997), and Sala-i-Martin et al. (2004). Unfortunately, similar robustness studies in which researchers examine the sensitivity of parameter estimates with respect to an extensive set of

---

5 Young and Holsteen (2015) find that out of 60 quantitative articles published in two of the top sociological journals, 85 percent contained at least one footnote (average 3.2) referencing additional robustness tests that were often unreported. Out of 164 footnotes, however, not a single one reported results that failed to support their findings.
additional, potentially influential controls “remain entirely absent in the literature on entrepreneurship” (Bjørnskov & Foss, 2016, p. 301).

The goal of this paper, then, was to conduct a robustness analysis for 44 possible determinants of opportunity- and necessity-motivated entrepreneurial activity while accounting for model uncertainty. To do this, we used a methodology proposed by Young and Holsteen (2015) and reported the robustness of the estimated coefficients to the choice of control variables from five different categories: (1) economic factors, (2) formal institutions, (3) cultural values, (4) legal origins, and (5) geographical conditions. Our goal was to assess whether the significance of the estimated coefficients depends on the set of control variables included in the model or if the results are consistent regardless of the assumptions made about the model. To that end, we present several robustness statistics, such as sign stability, significance rate, and a robustness ratio, that account not only for the uncertainty inherent in the data, as it is commonly practiced in classical statistical methods, but also for the uncertainty of the model assumptions, specifically the choice of control variables. In that sense, we focus on the weight of the evidence from thousands of estimations as opposed to a “single” or a few select estimates.

Our results can be summarized as follows. First, people who live in countries with higher levels of income inequality are less likely to engage in opportunity-motivated entrepreneurship (OME) and more likely to engage in entrepreneurial action out of necessity. Second, the level of economic development is only robustly correlated with necessity-motivated entrepreneurship (NME), but not with OME. Additionally, factor endowments such as physical and human capital do not appear to be robustly correlated with either type of entrepreneurship. Third, economic freedom is found to be the most robust determinant of both OME and NME across countries. Specifically, institutions associated with sound monetary policy, lower levels of corruption, and business and
investment freedoms have the strongest and most robust correlations with entrepreneurial action. Fourth, cultural values associated with Hofstede’s dimensions, social trust, ethnolinguistic fractionalization, and religious affiliation are not robustly correlated with OME or NME. Finally, countries with Scandinavian legal origins seem to have significantly higher levels of OME and lower levels of NME.

These findings have important implications for future cross-country entrepreneurship research. First, we identify the most robust predictors of opportunity- and necessity-motivated entrepreneurship. This set of robust variables, in turn, should become the starting point for future entrepreneurship research at the cross-country level. For example, our findings suggest that the most robust predictors of necessity-motivated entrepreneurship are net income inequality, economic development, economic freedom, and Scandinavian legal origins. Based on these findings, researchers who intend to model necessity-motivated entrepreneurship will want to include these variables in their estimations before considering other determinants of necessity entrepreneurship (or have a good theoretical reason why these variables should be excluded from their models). Second, we highlight the importance of variables such as income inequality and legal origins, which have so far received less attention in the entrepreneurship literature. While recent calls have emphasized the importance of studying income inequality with several recent contributions (e.g., Xavier-Oliveira, Laplume, & Pathak, 2015), the role of legal origins on entrepreneurial action is still significantly under researched and may prove a fruitful avenue for future work. Finally, while our findings provide an important validation test for some established theories (e.g., institutional theory), they also raise serious uncertainties about the direct role of cultural values and geography on entrepreneurial action. For instance, none of the measures of culture passed our robustness test, casting doubt on established cross-country relationships.
While evidence from our study reveals that economic and institutional factors and, to a lesser extent legal systems, demonstrate robust relationships (at some level) with NME and OME, confirmation of proposed connections with cultural values and geographic features proved entirely elusive. Geography has long been linked to important outcome variables, such as economic development and innovation (Gallup et al., 1999), but it has rarely been added as a control in cross-country entrepreneurship research. For this reason, we accept that its inclusion in our study is somewhat exploratory. The relationship between cultural values and entrepreneurship, on the other hand, has been quite extensively investigated (Terjesen et al., 2013), so our findings of a non-robust relationship deserve further contemplation.

The results of our study suggest that policies intended to spur different types of entrepreneurship might best be shaped only after examining the ways in which institutional, legal, and regulatory environments affect entrepreneurship in concert with culture, rather than looking for direct effects from cultural dimensions as is common in the empirical literature. This is consistent with recommendations from Hayton et al. (2002, p.49), who argue that “researchers should give greater attention to the interactions among cultural dimensions and the simultaneous influence of cultural, regulatory, and industry characteristics on aggregate entrepreneurship.” It is entirely possible that such an approach would be necessary to tease out the nuanced relationship between these two principal variables. And it follows that researchers might need to examine culture’s role in shaping entrepreneurship under various environmental conditions if they are to arrive at a clear (and robust) understanding of the connection.

Though this general research thrust has been relatively broad-ranging, the cultural dimension of greatest interest in past research has been Individualism-Collectivism (e.g, Baughn & Neupert, 2003; Mitchell et al., 2000), and this is perhaps with good reason. The theoretical fit is intuitively
sound: The hallmarks of individualist societies—e.g., personal control, autonomy, and incentive structures that reward individual accomplishment and creativity with social status—should naturally promote entrepreneurial propensities and the startup efforts that flow from these. Indeed, studies offer evidence of a causal relationship between individualism and entrepreneurial initiative (e.g., McGrath, MacMillan, & Scheinberg, 1992; Mueller & Thomas, 2001) and other related outcomes, including long-term economic growth and novel action (Gorodnichenko & Roland, 2011a, 2011b). However, Kyriacou (2016) argues convincingly that while individualist societies lean toward providing strong “incentives to invest, innovate, and accumulate wealth,” this influence is channeled, at least in part, through high quality political institutions that also result from the same forces. Empirically, once he accounts for the quality of political institutions, the significant relationship between cultural values and economic outcomes completely disappears. These findings suggest that the quality of institutions play an important mediating role in the relationship between culture and entrepreneurship.

In this respect, as Grief (2006) explains, collectivist societies tend to promote in-group interaction (usually defined by family, religion, tribe, or ethnic affiliation) and handle contract enforcement through cultural (informal) influences, whereas individualist societies encourage fluid transactional involvements across groups, backing up contracts by formal institutions (e.g., specialized systems, such as the courts). The former leans toward protecting entrenched commercial interests, making new ventures more difficult to launch and sustain. High quality governance systems, on the other hand, provide a more level playing field and open the door to new competition, including startups (Kyriacou, 2015). But the two are connected. In support of this claim, Nikolaev & Salahodjaev (2017) show that the origins of economic institutions are deeply rooted in the formation of cultural values over time. More specifically, using a two-stage
least squares analysis, their study finds strong causal evidence that the historical formation of personality traits, cultural attributes, and morality at the regional level shaped economic institutions such as competitive markets, law enforcement, and regulatory practices across countries and over time. This line of reasoning suggests that it is difficult to understand fully the influence of culture on entrepreneurship apart from first recognizing its impact on institutions, a view that would be consistent with our own findings.

Consequently, we do not mean to suggest that cultural values do not matter at all. Our findings do not test for alternative functional forms, such as non-linear and interactive relationships. It could be that the relationship between cultural values and entrepreneurship hinges on the level of economic development or its interaction with formal institutions, as we argue above. Indeed, previous research indicates that the relationship between formal institutions and culture (in general) can be far more complex and multi-directional than is modeled in this paper (Alesina & Giuliano, 2015). For instance, Minh & Hjortsø (2015) show that formal institutions influence not only small and medium-sized enterprise innovation, but also networking practices and social norms, which suggests that the relationship between culture and institutions can also run in the opposite direction. Similarly, previous studies find evidence that the relationship between culture, formal institutions, and income inequality is more complex than commonly assumed, which can lead to counter-intuitive and ambiguous empirical findings (e.g., Bennett & Nikolaev, 2016a, 2016b; Nikolaev, Boudreaux, & Salahodjaev, 2017).

Similarly, geography can also affect growth and innovation indirectly. A good example of this can be found in the settlements of European colonizers in the New World after 1500 (Acemoglu, Johnson, & Robinson, 2003, 2001; Diamond, 1999). Through its impact on outcomes such as crop yields and the spread of germs, geographical conditions consequently shaped the economic and
political institutions that took form to manage these challenges (Acemoglu et al., 2005). The prevalence of infectious diseases, more common to tropical areas, has also affected the emergence of cultural values, such as individualism-collectivism (Nikolaev et al. 2017), and can explain a significant share of current variations around the world related to economic development, human capital, and the propensity of societies to welcome and adopt new ideas (Thornhill & Fincher, 2014). Previous research also suggests a strong correlation between legal origins and the development of financial institutions, capital markets, government policies, and judicial institutions (e.g., La Porta, Lopez-de-Silanes, Shleifer, & Vishny, 1998). Thus, similar conclusions can be drawn with respect to geography and legal origins to that of cultural values.

In conclusion, our findings suggest that if culture, geography, or legal origins influence entrepreneurship, it is likely that these factors play a more nuanced role rather than a more direct role. Thus, researchers should look toward alternative functional forms, such as non-linear and interactive relationships and interpret previous findings with caution.

7. Limitations

As most empirical studies, ours also has important limitations. First, as Heckman (2005) observes, causal inference is impermanent in nature because it depends on prior model assumptions that can always change in the future. Robustness studies like ours help stimulate the process of knowledge creation by rigorously testing multiple models, allowing readers to focus on the weight of the evidence as opposed to an author’s preferred estimations. As Young and Holsteen (2015) maintain, however, the modeling space is “not only large, but also open ended—new additions to the model space can always be considered.” Robustness tests are about model transparency, which can reduce the problem of asymmetric information between researchers and
readers (Young, 2009). In that sense, the goal of this study is to provide analysis that is
developmental and compelling, but we also accept that it is not complete by any means.

One limitation of this study is that we are unable to test for alternative functional forms that
can prove helpful in explaining the underlying mechanisms of the proposed relationships. In this
respect, our conclusion is not necessarily that cultural values, geography, or legal origins do not
matter. We offer this caveat in recognition of the fact that the relationship between these and other
variables can be far more complex than we are able to capture with our model (Wennberg et al.,
2013). Culture may function as a moderator or a mediator and work through a variety of channels,
such as formal institutions and economic growth, that can in turn affect cultural values (e.g.,
Dennis, 2011). Computationally, it is very difficult to account for all of these possibilities. To wit,
including the moderating effects of 44 variables would increase exponentially the number of
models that would need to be tested. Therefore, we only focus on linear relationships in this paper;
and even at that, our analysis is already very demanding computationally. Nonetheless, our
findings are important because they raise crucial questions about findings from prior research that
are based on linear methods similar to ours. This emphasizes the importance of more robust testing
in the future.

Another limitation of our findings is that they remain silent about the direction of causality. In
the absence of a randomized controlled experiment, which is infeasible for this type of analysis, it
is always challenging to establish causal relationships. It is possible, for instance, that income
inequality, one of our robust predictors of NME, is as much a cause of entrepreneurial action as it
is a product of a highly dynamic entrepreneurial environment. Inequality motivates entrepreneurs
to take greater risks and enter new markets that generate substantial wealth for themselves, but this
can lead to an even wider gap between the rich and the poor. On the other hand, higher levels of
income inequality may stifle income mobility and create barriers for entrepreneurs to enter and exploit opportunities, leading to lower rates of OME. Similarly, previous studies suggest that institutions not only can influence entrepreneurial outcomes, but entrepreneurs may also influence institutional climate, especially in challenging environments such as emerging market and transitioning economies (Welter & Smallbone, 2011). We are unable to test such dynamic processes with cross-sectional data.

Similarly, we are unable to address issues associated with unobserved heterogeneity—i.e., potential bias from omitting important variables that can be correlated with both independent and dependent variables, leading to spurious correlations. Future studies will have to address these issues using instrumental variables (IV), Heckman selection estimators, fixed-effects models, or difference-in-difference estimators. IV analysis, for instance, can be used to mitigate the issue of omitted variable bias, and in some instances reverse causality, but under the strong assumptions of instrument relevance and exogeneity (the so called exclusion restriction). This creates second-order questions of uncertainty about how well the instrument meets these conditions (Hahn & Hausman, 2003). As Glaser (2008) also notes, more sophisticated technical models are often less transparent to readers and provide researchers with even greater degrees of discretion to report non-robust findings. While such models will greatly enrich the model space, they also make the need for robustness testing even more imperative.

Finally, definitions of entrepreneurship have varied significantly in the literature—from self-employment and innovation rates to nascent entrepreneurship and total early-stage entrepreneurs. In this paper, we focus on only two possible measures of entrepreneurship—OME and NME. Indeed, there are others. But hopefully our analysis can easily be incorporated in future studies that use different definitions of entrepreneurship.
The current results, then, should be viewed as a call for more robustness studies in the field of entrepreneurship. As we show in this paper, there is significant heterogeneity in how entrepreneurship scholars select their models. But the global blossoming of entrepreneurship and the availability of more international data presents one of the most exciting opportunities for entrepreneurship scholars to study how different economic, institutional, cultural, historical, and political factors affect entrepreneurship (Terjesen, Hessels, & Li, 2016). More importantly, the analysis presented in this paper can easily be incorporated with standard statistical packages such as Stata (Young and Holsteen, 2015). This would allow future studies to perform robustness analyses that can make empirical findings more compelling and less prone to non-robust, trivial, and false-positive results.

References


## Appendix

### Table 1: Heterogeneity in the Empirical Literature

<table>
<thead>
<tr>
<th>Paper</th>
<th>Year</th>
<th>Estimation</th>
<th>Variables</th>
<th>N</th>
<th>Dep Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wong et al.</td>
<td>2005</td>
<td>OLS</td>
<td>Log GDP (* with OME, NME, and TEA), Growth in capital investment (+, * with OME, NME, and TEA)</td>
<td>32</td>
<td>OME, NME, TEA</td>
</tr>
<tr>
<td>Klapper et al.</td>
<td>2006</td>
<td>OLS</td>
<td>Industry share (-, non), industry-level new entry ratios (-, non) and entry costs interaction (-, *)</td>
<td>23</td>
<td>Entry Rates</td>
</tr>
<tr>
<td>Freytag and Thurik</td>
<td>2007</td>
<td>OLS</td>
<td>EF: Regulation (-, non), Post-Communist (+, non), Life Expectancy (-, non), Health (-, non)</td>
<td>27</td>
<td>Self-Employment (actual)</td>
</tr>
<tr>
<td>Wennekers et al.</td>
<td>2007</td>
<td>OLS</td>
<td>Uncertainty Avoidance (+/<em>), GDP (+/</em>), Gini (+<em>), Replacement Unemployment (-, <em>, non), Female Labor Share (- (</em>, non), Share Services (+,</em>,non)</td>
<td>63</td>
<td>Business ownership rate</td>
</tr>
<tr>
<td>Uhlner and Thurik</td>
<td>2007</td>
<td>OLS</td>
<td>Post-materialism (-, *, non), GDP per capita (-, *, non), GDP squared (+, *, non), Education - secondary (-, *, non), Education-tertiary(+, *, non), Life Satisfaction</td>
<td>27</td>
<td>TEA</td>
</tr>
<tr>
<td>Sobel et al.</td>
<td>2007</td>
<td>OLS</td>
<td>Economic freedom index, average tariff rate (-, ), domestic entry barriers (-), percent male (+, non), median age (-, *), GDP (+, non), unemployment (-, non), domestic credit availability (-,non), foreign capital (-, non), political stability (-/+,non)</td>
<td>21</td>
<td>TEA</td>
</tr>
<tr>
<td>Nyström</td>
<td>2008</td>
<td>Panel</td>
<td>GDP (-, sig), Unemployment (-, *), Gov Size (+, *), Rule of Law (+, *), Sound Money (-, non), International Trade (-, non), Regulation (+, *)</td>
<td>23</td>
<td>Self-Employment (actual)</td>
</tr>
<tr>
<td>McMullen et al.</td>
<td>2008</td>
<td>OLS</td>
<td>GDP (-, * with both OME, NME), Rule of Law (-, * with OME, non with NME), Labor Freedom (-, * with both OME and NME), Monetary Freedom (-, * with NME), Fiscal Freedom (-, * with NME), insignificant results for the rest freedom sub-indexes</td>
<td>37</td>
<td>OME, NME</td>
</tr>
<tr>
<td>Bjørnskov and Foss</td>
<td>2008</td>
<td>OLS</td>
<td>GDP (-, * with NME and non with OME), post-communist (+/-, non), education, income inequality, employment, exchange rate volatility, investment price level, market capitalization (used for robustness)</td>
<td>29</td>
<td>OME, NME, TEA</td>
</tr>
<tr>
<td>Powell and Rodet</td>
<td>2012</td>
<td>OLS</td>
<td>Cultural Index (+,*), GDP (-, non), Gov Size (+, non), Sound Money (+, non), Rule of Law (-, non), Regulation (+/-, non)</td>
<td>21</td>
<td>TEA</td>
</tr>
<tr>
<td>Autio et al.</td>
<td>2013</td>
<td>Multi-level</td>
<td>GDP (+, *, non), Population (+, non), Collectivism (+, *,non), Assertiveness (+, *, non), Uncertainty Avoidance (+, *,non), Performance Orientation (+, *, non)</td>
<td>42</td>
<td>Entry E-ship, Growth Aspirations</td>
</tr>
</tbody>
</table>
Table 2: Description of Variables and Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Entrepreneurship</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NME</td>
<td>Necessity-Motivated Entrepreneurship. Shows percentage of total early stage</td>
<td>GEM (2016)</td>
<td>73</td>
<td>24.90</td>
<td>9.78</td>
<td>5.70</td>
<td>46.84</td>
</tr>
<tr>
<td></td>
<td>entrepreneurs who are involved in entrepreneurship because they have no other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>options for work, average over the period 2001-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OME</td>
<td>Opportunity-Motivated Entrepreneurship. Shows percentage of total early stage</td>
<td>GEM (2016)</td>
<td>73</td>
<td>48.88</td>
<td>10.88</td>
<td>29</td>
<td>78.94</td>
</tr>
<tr>
<td></td>
<td>entrepreneurs who are involved in entrepreneurship because (1) they are</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>driven by opportunity or (2) claim that their main driver for being involved</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>in entrepreneurship is being independent or increasing their income, average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>over the period 2001-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Area 1: Economic Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log GDP</td>
<td>Log of real GDP per capita, PPP (2011US$), average over the period 2001-2015</td>
<td>Penn World Tables (2016)</td>
<td>72</td>
<td>9.73</td>
<td>0.98</td>
<td>6.66</td>
<td>11.42</td>
</tr>
<tr>
<td>Human Capital</td>
<td>Human capital index, based on years of schooling and returns to education,</td>
<td>Penn World Tables (2016)</td>
<td>73</td>
<td>2.65</td>
<td>0.64</td>
<td>1.12</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td>average over the period 2001-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Av Hours Worked</td>
<td>Average annual hours worked by persons engaged, average over the period</td>
<td>Penn World Tables (2016)</td>
<td>56</td>
<td>1890</td>
<td>249</td>
<td>1431</td>
<td>2362</td>
</tr>
<tr>
<td></td>
<td>2001-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiteracy Rate</td>
<td>Adult illiteracy rate, 25-64 years, both sexes, average 2001-2015</td>
<td>UNESCO (2016)</td>
<td>53</td>
<td>19.15</td>
<td>20.55</td>
<td>0.25</td>
<td>82.96</td>
</tr>
<tr>
<td>Capital Stock</td>
<td>Capital stock at current PPPs (in mil. 2011US$), average over the period</td>
<td>Penn World Tables (2016)</td>
<td>73</td>
<td>2745</td>
<td>6218</td>
<td>33</td>
<td>42300</td>
</tr>
<tr>
<td></td>
<td>2001-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solt (net)</td>
<td>Gini coefficient representing relative net income inequality. Average over</td>
<td>Solt (2016)</td>
<td>73</td>
<td>38.03</td>
<td>9.96</td>
<td>21.94</td>
<td>62.77</td>
</tr>
<tr>
<td></td>
<td>period 2001–2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solt (gross)</td>
<td>Gini coefficient representing relative gross income inequality. Average</td>
<td>Solt (2016)</td>
<td>73</td>
<td>44.25</td>
<td>6.90</td>
<td>30.94</td>
<td>66.23</td>
</tr>
<tr>
<td></td>
<td>over period 2001–2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td>Number of persons engaged (in millions), average over the period 2001-2015</td>
<td>Penn World Tables (2016)</td>
<td>73</td>
<td>33.18</td>
<td>101.67</td>
<td>0.16</td>
<td>745.69</td>
</tr>
<tr>
<td>Exchange Rate</td>
<td>Exchange rate, national currency/USD (market+estimated), average over the</td>
<td>Penn World Tables (2016)</td>
<td>73</td>
<td>524.81</td>
<td>2318.38</td>
<td>0.61</td>
<td>15987.5</td>
</tr>
<tr>
<td></td>
<td>period 2001-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
<td>----</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
</tr>
</tbody>
</table>

**Area 2: Formal Institutions**

**EF Index**

Index of economic freedom evaluates countries on broad dimensions of economic environment over which governments typically exercise policy control. Values range from 0 (least free) to 100 (most free).

**EF: Corruption**

Sub index of economic freedom which measures the extent to which corruption prevails in a country.

**EF: Fiscal**

The fiscal freedom component is a composite measure of the burden of taxes that reflects both marginal tax rates and the overall level of taxation, including direct and indirect taxes imposed by all levels of government.

**EF: Government**

Government Spending sub-component captures burden of government expenditures, which include consumption by state and all transfer payments related to various entitlement programs.

**EF: Business**

Sub index of economic freedom which measures the extent to which the regulatory and infrastructure environments constrain the efficient operation of businesses.

**EF: Labor**

Sub index of economic freedom that takes into account various aspects of the legal and regulatory framework of a country's labor market.

**EF: Monetary**

Sub index of economic freedom that measures sound monetary policy and price stability.

**EF: Trade**

Trade freedom is a sub-index of economic freedom that measures of the extent of tariff and non-tariff barriers that have impact on bilateral trade.

**EF: Investment**

Sub index of economic freedom that evaluates a variety of regulatory restrictions that typically are imposed on investment.

**EF: Financial**

Sub index of economic freedom that measures banking system efficiency.

**Democracy**

Index that is measured as average of civil rights and political liberties.

**Area 3: Cultural Values**

**Individualism**

Individualism index that measures the degree to which a society accepts and reinforces individualist or collectivist values. The index ranges from 0 (most collectivistic) and 100 (most individualistic).

**Power Dist**

Power Distance index that measures the extent to which the less powerful members of organizations and institutions (in society and
the family) accept and expect that power is distributed unequally, with higher values reflecting higher acceptance of inequality in the distribution of power by those below.

| Masculinity | Masculinity index that reflects dominance of men over women and to the dominance of "male" values such as assertiveness and competitiveness versus norms of caring and modesty | Hofstede et al. (2010) | 73 | 49.51 | 18.94 | 5 | 100 |
| Uncertainty | Uncertainty Avoidance index that refers to society’s tolerance for uncertainty and the extent to which members of society feel either uncomfortable or comfortable in situations that are novel, unknown, surprising, different from usual. | Hofstede et al. (2010) | 73 | 64.79 | 22.14 | 8 | 100 |
| Long term | Long Term Orientation index that measures the extent to which societies maintain links with their own past while dealing with the challenges of the present and future. Societies who score low on this index prefer to maintain timehonored traditions and norms while viewing societal changes with suspicion. | Hofstede et al. (2010) | 66 | 42.98 | 22.26 | 4 | 88 |
| Indulgence | Indulgence index that measures the degree to which societies allow free gratification of basic and natural human drives related to enjoying life and having fun, with higher values of this index reflecting more indulgent societies with less restrictive social norms. | Hofstede et al. (2010) | 63 | 48.83 | 22.86 | 0 | 100 |
| Ethno | Index that captures the probability that two individuals, selected at random from a country's population, will belong to different ethnic groups. | Alesina et al. (2003) | 73 | 0.39 | 0.23 | 0.01 | 0.85 |
| Trust | Measures the proportion of people who say they "most people can be trusted," average over period 1999-2014. | World Values Survey (2016) | 59 | 0.27 | 0.15 | 0.04 | 0.70 |
| Religion (base=Other) | Share of Muslim population | La Porta et al. (2000) | 73 | 13.61 | 28.40 | 0 | 99.40 |
| Catholic | Share of Catholic population | La Porta et al. (2000) | 73 | 39.84 | 39.26 | 0 | 96.90 |
| Protestant | Share of Protestant population | La Porta et al. (2000) | 73 | 15.93 | 26.10 | 0 | 97.80 |

**Area 4: Geography & Legal Origins**

**Legal Origins (base = German)**

| Legal Origins | Dummy = 1 if legal origins are Great Britain, 0 = otherwise | La Porta et al. (2000) | 73 | 0.29 | 0.46 | 0 | 1 |
| Legor: French | Dummy = 1 if legal origins are French, 0 = otherwise | La Porta et al. (2000) | 73 | 0.41 | 0.50 | 0 | 1 |
| Legor: Scand | Dummy = 1 if legal origins are Scandinavian, 0 = otherwise | La Porta et al. (2000) | 73 | 0.07 | 0.25 | 0 | 1 |
| Legor: Socialist | Dummy = 1 if legal origins are Socialist, 0 = otherwise | La Porta et al. (2000) | 73 | 0.18 | 0.39 | 0 | 1 |

**Geography**

| Latitude | Value of the latitude of a country’s approximate geodesic centroid | La Porta et al. (2000) | 73 | 24.73 | 28.19 | -41.8 | 65.0 |
| Tropics | Proportion of land area located in tropical region | Gallup et al., 1999. | 73 | 0.38 | 0.47 | 0 | 1 |

**Continental Dummies (base = North America)**

| Cont: Africa | Dummy = 1 if continent is Africa, 0 = otherwise | Ashraf & Galor (2013) | 73 | 0.15 | 0.36 | 0 | 1 |
| Cont: Asia | Dummy = 1 if continent is Asia, 0 = otherwise | Ashraf & Galor (2013) | 73 | 0.19 | 0.40 | 0 | 1 |
| Cont: Europe | Dummy = 1 if continent is Europe, 0 = otherwise | Ashraf & Galor (2013) | 73 | 0.38 | 0.49 | 0 | 1 |
| Cont: Oceania | Dummy = 1 if continent is Oceania, 0 = otherwise | Ashraf & Galor (2013) | 73 | 0.03 | 0.16 | 0 | 1 |
| Cont: S America | Dummy = 1 if continent is South America, 0 = otherwise | Ashraf & Galor (2013) | 73 | 0.11 | 0.31 | 0 | 1 |

**Notes:** Authors’ calculations.
### Table 3: Economic Controls

<table>
<thead>
<tr>
<th>Variable</th>
<th>$(\beta)$</th>
<th>R Ratio</th>
<th>Sign Stab</th>
<th>Sig Rate</th>
<th>+</th>
<th>+ &amp; sig</th>
<th>-</th>
<th>- &amp; sig</th>
<th>N</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Opportunity-Motivated Entrepreneurship</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solt (net)</td>
<td>-0.5499</td>
<td>-1.8367</td>
<td>100</td>
<td>73</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
<td>73</td>
<td>weakly robust</td>
</tr>
<tr>
<td>Solt (gross)</td>
<td>-0.4436</td>
<td>-1.8279</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>50</td>
<td>73</td>
<td>weakly robust</td>
</tr>
<tr>
<td>Exchange Rate</td>
<td>0.0007</td>
<td>1.3977</td>
<td>100</td>
<td>41</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Service</td>
<td>-0.4073</td>
<td>-1.3213</td>
<td>97</td>
<td>38</td>
<td>3</td>
<td>0</td>
<td>97</td>
<td>38</td>
<td>72</td>
<td>not robust</td>
</tr>
<tr>
<td>Employment</td>
<td>-0.0136</td>
<td>-1.1151</td>
<td>97</td>
<td>20</td>
<td>3</td>
<td>0</td>
<td>97</td>
<td>20</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Log GDP</td>
<td>1.6668</td>
<td>0.4950</td>
<td>76</td>
<td>10</td>
<td>76</td>
<td>10</td>
<td>76</td>
<td>0</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Illiteracy Rate</td>
<td>0.0724</td>
<td>0.4469</td>
<td>77</td>
<td>1</td>
<td>77</td>
<td>1</td>
<td>23</td>
<td>0</td>
<td>53</td>
<td>not robust</td>
</tr>
<tr>
<td>Capital Stock</td>
<td>0.0000</td>
<td>0.4301</td>
<td>73</td>
<td>3</td>
<td>73</td>
<td>3</td>
<td>73</td>
<td>0</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Human Capital</td>
<td>1.4970</td>
<td>0.2989</td>
<td>68</td>
<td>5</td>
<td>68</td>
<td>5</td>
<td>32</td>
<td>0</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Industry</td>
<td>-0.0433</td>
<td>-0.2366</td>
<td>68</td>
<td>0</td>
<td>32</td>
<td>0</td>
<td>68</td>
<td>0</td>
<td>72</td>
<td>not robust</td>
</tr>
<tr>
<td>Av Hours Worked</td>
<td>0.0011</td>
<td>0.1373</td>
<td>63</td>
<td>0</td>
<td>63</td>
<td>0</td>
<td>37</td>
<td>0</td>
<td>56</td>
<td>not robust</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Necessity-Motivated Entrepreneurship</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solt (net)</td>
<td>0.3613</td>
<td>2.0429</td>
<td>100</td>
<td>70</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>73</td>
<td>robust</td>
</tr>
<tr>
<td>Log GDP</td>
<td>-3.7898</td>
<td>-1.4597</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>50</td>
<td>73</td>
<td>weakly robust</td>
</tr>
<tr>
<td>Solt (gross)</td>
<td>0.1840</td>
<td>0.9705</td>
<td>99</td>
<td>7</td>
<td>99</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Exchange Rate</td>
<td>-0.0047</td>
<td>-0.9188</td>
<td>93</td>
<td>21</td>
<td>7</td>
<td>0</td>
<td>93</td>
<td>21</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Service</td>
<td>0.2207</td>
<td>0.8451</td>
<td>89</td>
<td>26</td>
<td>89</td>
<td>25</td>
<td>11</td>
<td>0</td>
<td>72</td>
<td>not robust</td>
</tr>
<tr>
<td>Employment</td>
<td>0.0076</td>
<td>0.7438</td>
<td>90</td>
<td>15</td>
<td>90</td>
<td>15</td>
<td>10</td>
<td>0</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Capital Stock</td>
<td>0.0000</td>
<td>0.4852</td>
<td>83</td>
<td>1</td>
<td>83</td>
<td>1</td>
<td>17</td>
<td>1</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Illiteracy Rate</td>
<td>-0.0581</td>
<td>-0.3738</td>
<td>59</td>
<td>17</td>
<td>41</td>
<td>1</td>
<td>59</td>
<td>17</td>
<td>53</td>
<td>not robust</td>
</tr>
<tr>
<td>Human Capital</td>
<td>-1.1501</td>
<td>-0.2404</td>
<td>62</td>
<td>20</td>
<td>38</td>
<td>5</td>
<td>62</td>
<td>15</td>
<td>73</td>
<td>not robust</td>
</tr>
<tr>
<td>Av Hours Worked</td>
<td>0.0009</td>
<td>0.1256</td>
<td>53</td>
<td>1</td>
<td>53</td>
<td>1</td>
<td>47</td>
<td>0</td>
<td>56</td>
<td>not robust</td>
</tr>
<tr>
<td>Industry</td>
<td>-0.0130</td>
<td>-0.0870</td>
<td>53</td>
<td>10</td>
<td>53</td>
<td>1</td>
<td>47</td>
<td>9</td>
<td>72</td>
<td>not robust</td>
</tr>
</tbody>
</table>

**Notes:** We follow the methodology outlined by Young and Holsteen (2015). Results are summary of the modelling distribution for each variable of interest (e.g., Log GDP) based on 8,129 unique combinations of the 14 core variables described in section 2. Since the procedure is computationally very intensive, we treat geographic, religious, and legal origins dummies as vectors of variables that enter the estimations together (e.g., all religious dummies enter the regression together instead of individually).

$(\beta)$ = the average $\beta$ coefficient across all 8,129 estimations.

R Ratio = Robustness Ratio. If higher than 2, it suggests robustness (Young and Holsteen, 2015).

+ = % of models in which the variable enters with a positive sign.

+ & sig = % of models in which the variable enters with a positive & significant sign.

- = % of models in which the variable enters with a negative sign.

- & sig = % of models in which the variable enters with a negative & significant sign.

sig stab = sign stability indicating the percentage of models that have the same sign.

sig rate = significance rate indicating the percentage of models that report statistically significant coefficient. A significance rate of 95% or higher indicates “strong” robustness while a significance rate of 50% sets a lower bound for “weak” robustness (Raftery, 1995).

N = number of observations
Table 4: Formal Institutions

<table>
<thead>
<tr>
<th>Variable</th>
<th>(β)</th>
<th>R Ratio</th>
<th>Sign Stab</th>
<th>Sign Rate</th>
<th>+ &amp; sig</th>
<th>- &amp; sig</th>
<th>N</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opportunity-Motivated Entrepreneurship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF: Monetary</td>
<td>0.58</td>
<td>4.13</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF: Corruption</td>
<td>0.39</td>
<td>4.02</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF Index</td>
<td>0.66</td>
<td>3.68</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF: Business</td>
<td>0.55</td>
<td>3.28</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF: Labor</td>
<td>0.14</td>
<td>1.44</td>
<td>100</td>
<td>26</td>
<td>100</td>
<td>26</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF: Trade</td>
<td>0.25</td>
<td>1.43</td>
<td>100</td>
<td>24</td>
<td>100</td>
<td>24</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF: Financial</td>
<td>0.16</td>
<td>1.40</td>
<td>100</td>
<td>22</td>
<td>100</td>
<td>22</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF: Fiscal</td>
<td>-0.20</td>
<td>-1.27</td>
<td>95</td>
<td>38</td>
<td>5</td>
<td>0</td>
<td>95</td>
<td>38</td>
</tr>
<tr>
<td>EF: Government</td>
<td>0.10</td>
<td>0.73</td>
<td>83</td>
<td>24</td>
<td>83</td>
<td>23</td>
<td>17</td>
<td>73</td>
</tr>
<tr>
<td>EF: Investment</td>
<td>0.05</td>
<td>0.43</td>
<td>87</td>
<td>0</td>
<td>87</td>
<td>0</td>
<td>13</td>
<td>73</td>
</tr>
<tr>
<td>Democracy</td>
<td>-0.50</td>
<td>-0.29</td>
<td>68</td>
<td>4</td>
<td>32</td>
<td>0</td>
<td>68</td>
<td>4</td>
</tr>
<tr>
<td>Necessity-Motivated Entrepreneurship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF Index</td>
<td>-0.54</td>
<td>-3.92</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Corruption</td>
<td>-0.28</td>
<td>-3.85</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Monetary</td>
<td>-0.39</td>
<td>-3.37</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Business</td>
<td>-0.34</td>
<td>-2.49</td>
<td>100</td>
<td>82</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Investment</td>
<td>-0.16</td>
<td>-1.82</td>
<td>100</td>
<td>52</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Fiscal</td>
<td>0.22</td>
<td>1.74</td>
<td>100</td>
<td>66</td>
<td>100</td>
<td>66</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>EF: Labor</td>
<td>-0.12</td>
<td>-1.68</td>
<td>100</td>
<td>47</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Financial</td>
<td>-0.13</td>
<td>-1.44</td>
<td>100</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Trade</td>
<td>-0.25</td>
<td>-1.42</td>
<td>100</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>EF: Government</td>
<td>-0.07</td>
<td>-0.68</td>
<td>79</td>
<td>27</td>
<td>21</td>
<td>2</td>
<td>79</td>
<td>73</td>
</tr>
<tr>
<td>Democracy</td>
<td>-0.41</td>
<td>-0.33</td>
<td>66</td>
<td>6</td>
<td>34</td>
<td>0</td>
<td>66</td>
<td>73</td>
</tr>
</tbody>
</table>

Notes: We follow the methodology outlined by Young and Holsteen (2015). Results are summary of the modelling distribution for each variable of interest (e.g., Log GDP) based on 8,129 unique combinations of the 14 core variables described in section 2. Since the procedure is computationally very intensive, we treat geographic, religious, and legal origins dummies as vectors of variables that enter the estimations together (e.g., all religious dummies enter the regression together instead of individually).

(β) = the average β coefficient across all 8,129 estimations.
R Ratio = Robustness Ratio. If higher than 2, it suggests robustness (Young and Holsteen, 2015).
+ = % of models in which the variable enters with a positive sign.
+ & sig = % of models in which the variable enters with a positive & significant sign.
- = % of models in which the variable enters with a negative sign.
- & sig = % of models in which the variable enters with a negative & significant sign.
sign stab = sign stability indicating the percentage of models that have the same sign.
sig rate = significance rate indicating the percentage of models that report statistically significant coefficient. A significance rate of 95% or higher indicates "strong" robustness while a significance rate of 50% sets a lower bound for "weak" robustness (Raftery, 1995).
N = number of observations
Table 5: Culture

<table>
<thead>
<tr>
<th>Variable</th>
<th>R Ratio</th>
<th>Sign Stab</th>
<th>Sig Rate</th>
<th>+</th>
<th>+ &amp; sig</th>
<th>-</th>
<th>- &amp; sig</th>
<th>N</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Opportunity-Motivated Entrepreneurship</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indulgence</td>
<td>0.12</td>
<td>1.55</td>
<td>100</td>
<td>41</td>
<td>100</td>
<td>41</td>
<td>0</td>
<td>0</td>
<td>63</td>
</tr>
<tr>
<td>Masculinity</td>
<td>-0.09</td>
<td>-1.26</td>
<td>100</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>24</td>
<td>73</td>
</tr>
<tr>
<td>Trust</td>
<td>17.81</td>
<td>1.22</td>
<td>100</td>
<td>24</td>
<td>100</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td>Catholic</td>
<td>-0.05</td>
<td>-0.90</td>
<td>98</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>98</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Ethno</td>
<td>4.89</td>
<td>0.70</td>
<td>89</td>
<td>0</td>
<td>89</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Muslim</td>
<td>-0.04</td>
<td>-0.53</td>
<td>93</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>-0.04</td>
<td>-0.48</td>
<td>78</td>
<td>7</td>
<td>22</td>
<td>0</td>
<td>78</td>
<td>7</td>
<td>73</td>
</tr>
<tr>
<td>Power Dist</td>
<td>-0.04</td>
<td>-0.43</td>
<td>73</td>
<td>6</td>
<td>27</td>
<td>0</td>
<td>73</td>
<td>6</td>
<td>73</td>
</tr>
<tr>
<td>Individualism</td>
<td>0.04</td>
<td>0.42</td>
<td>77</td>
<td>2</td>
<td>77</td>
<td>2</td>
<td>23</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Long term</td>
<td>-0.01</td>
<td>-0.13</td>
<td>62</td>
<td>3</td>
<td>38</td>
<td>0</td>
<td>62</td>
<td>3</td>
<td>66</td>
</tr>
<tr>
<td>Protestant</td>
<td>-0.01</td>
<td>-0.06</td>
<td>53</td>
<td>5</td>
<td>53</td>
<td>3</td>
<td>47</td>
<td>2</td>
<td>73</td>
</tr>
<tr>
<td><strong>Necessity-Motivated Entrepreneurship</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masculinity</td>
<td>0.08</td>
<td>1.52</td>
<td>100</td>
<td>36</td>
<td>100</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Catholic</td>
<td>0.06</td>
<td>1.35</td>
<td>99</td>
<td>23</td>
<td>99</td>
<td>23</td>
<td>1</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Trust</td>
<td>-13.18</td>
<td>-1.19</td>
<td>99</td>
<td>32</td>
<td>1</td>
<td>0</td>
<td>99</td>
<td>32</td>
<td>65</td>
</tr>
<tr>
<td>Ethno</td>
<td>-6.56</td>
<td>-1.17</td>
<td>97</td>
<td>21</td>
<td>3</td>
<td>0</td>
<td>97</td>
<td>21</td>
<td>73</td>
</tr>
<tr>
<td>Power Dist</td>
<td>0.06</td>
<td>0.84</td>
<td>90</td>
<td>16</td>
<td>90</td>
<td>16</td>
<td>10</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>0.05</td>
<td>0.79</td>
<td>92</td>
<td>8</td>
<td>92</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Indulgence</td>
<td>-0.05</td>
<td>-0.75</td>
<td>94</td>
<td>16</td>
<td>6</td>
<td>0</td>
<td>94</td>
<td>16</td>
<td>63</td>
</tr>
<tr>
<td>Muslim</td>
<td>0.04</td>
<td>0.59</td>
<td>84</td>
<td>3</td>
<td>84</td>
<td>3</td>
<td>16</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Individualism</td>
<td>-0.04</td>
<td>-0.50</td>
<td>79</td>
<td>6</td>
<td>21</td>
<td>0</td>
<td>79</td>
<td>6</td>
<td>73</td>
</tr>
<tr>
<td>Protestant</td>
<td>0.03</td>
<td>0.37</td>
<td>66</td>
<td>4</td>
<td>66</td>
<td>3</td>
<td>34</td>
<td>1</td>
<td>73</td>
</tr>
<tr>
<td>Long term</td>
<td>-0.01</td>
<td>-0.11</td>
<td>51</td>
<td>1</td>
<td>49</td>
<td>0</td>
<td>51</td>
<td>1</td>
<td>66</td>
</tr>
</tbody>
</table>

Notes: We follow the methodology outlined by Young and Holsteen (2015). Results are summary of the modelling distribution for each variable of interest (e.g., Log GDP) based on 8,129 unique combinations of the 14 core variables described in section 2. Since the procedure is computationally very intensive, we treat geographic, religious, and legal origins dummies as vectors of variables that enter the estimations together (e.g., all religious dummies enter the regression together instead of individually). Religion: Others used as a base.

(\( \beta \)) = the average \( \beta \) coefficient across all 8,129 estimations.

R Ratio = Robustness Ratio. If higher than 2, it suggests robustness (Young and Holsteen, 2015).

+ = % of models in which the variable enters with a positive sign.

+ & sig = % of models in which the variable enters with a positive & significant sign.

- = % of models in which the variable enters with a negative sign.

- & sig = % of models in which the variable enters with a negative & significant sign.

Sign stab = sign stability indicating the percentage of models that have the same sign.

Sign rate = significance rate indicating the percentage of models that report statistically significant coefficient. A significance rate of 95% or higher indicates “strong” robustness while a significance rate of 50% sets a lower bound for “weak” robustness (Raftery, 1995).

N = number of observations.
Table 6: Legal Origins & Geography

<table>
<thead>
<tr>
<th>Variables</th>
<th>(β)</th>
<th>R Ratio</th>
<th>Sign Stab</th>
<th>Sign Rate</th>
<th>+ &amp; sig</th>
<th>- &amp; sig</th>
<th>N</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opportunity-Motivated Entrepreneurship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legor: Scand</td>
<td>11.59</td>
<td>1.77</td>
<td>100</td>
<td>66</td>
<td>100</td>
<td>66</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cont: Oceania</td>
<td>18.07</td>
<td>1.44</td>
<td>100</td>
<td>47</td>
<td>100</td>
<td>47</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tropics</td>
<td>5.54</td>
<td>1.25</td>
<td>100</td>
<td>11</td>
<td>100</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cont: Asia</td>
<td>4.81</td>
<td>0.92</td>
<td>98</td>
<td>4</td>
<td>98</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Legor: Socialist</td>
<td>-5.09</td>
<td>-0.87</td>
<td>90</td>
<td>16</td>
<td>10</td>
<td>0</td>
<td>90</td>
<td>16</td>
</tr>
<tr>
<td>Cont: Europe</td>
<td>-4.29</td>
<td>-0.85</td>
<td>98</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>98</td>
<td>0</td>
</tr>
<tr>
<td>Cont: S America</td>
<td>4.02</td>
<td>0.72</td>
<td>92</td>
<td>4</td>
<td>92</td>
<td>4</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Cont: Africa</td>
<td>3.65</td>
<td>0.53</td>
<td>86</td>
<td>2</td>
<td>86</td>
<td>2</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Latitude</td>
<td>0.05</td>
<td>0.38</td>
<td>63</td>
<td>13</td>
<td>63</td>
<td>13</td>
<td>37</td>
<td>0</td>
</tr>
<tr>
<td>Legor: French</td>
<td>-1.66</td>
<td>-0.32</td>
<td>70</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>Legor: GB</td>
<td>-1.63</td>
<td>-0.32</td>
<td>73</td>
<td>0</td>
<td>27</td>
<td>0</td>
<td>73</td>
<td>0</td>
</tr>
<tr>
<td>Necessity-Motivated Entrepreneurship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legor: Scand</td>
<td>-9.05</td>
<td>-1.40</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Tropics</td>
<td>-5.27</td>
<td>-1.32</td>
<td>100</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>24</td>
</tr>
<tr>
<td>Legor: Socialist</td>
<td>5.89</td>
<td>1.08</td>
<td>100</td>
<td>8</td>
<td>100</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cont: Europe</td>
<td>-2.90</td>
<td>-0.70</td>
<td>90</td>
<td>4</td>
<td>10</td>
<td>0</td>
<td>90</td>
<td>4</td>
</tr>
<tr>
<td>Cont: Oceania</td>
<td>-5.27</td>
<td>-0.69</td>
<td>88</td>
<td>16</td>
<td>12</td>
<td>0</td>
<td>88</td>
<td>16</td>
</tr>
<tr>
<td>Cont: S America</td>
<td>1.77</td>
<td>0.39</td>
<td>75</td>
<td>1</td>
<td>75</td>
<td>1</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>Legor: French</td>
<td>-1.77</td>
<td>-0.34</td>
<td>79</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>79</td>
<td>0</td>
</tr>
<tr>
<td>Cont: Africa</td>
<td>1.68</td>
<td>0.27</td>
<td>65</td>
<td>2</td>
<td>65</td>
<td>2</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>Latitude</td>
<td>-0.01</td>
<td>-0.14</td>
<td>60</td>
<td>3</td>
<td>40</td>
<td>0</td>
<td>60</td>
<td>3</td>
</tr>
<tr>
<td>Legor: GB</td>
<td>-0.40</td>
<td>-0.09</td>
<td>61</td>
<td>0</td>
<td>39</td>
<td>0</td>
<td>61</td>
<td>0</td>
</tr>
<tr>
<td>Cont: Asia</td>
<td>-0.40</td>
<td>-0.09</td>
<td>61</td>
<td>0</td>
<td>39</td>
<td>0</td>
<td>61</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes: We follow the methodology outlined by Young and Holsteen (2015). Results are summary of the modelling distribution for each variable of interest (e.g., Log GDP) based on 8,129 unique combinations of the 14 core variables described in section 2. Since the procedure is computationally very intensive, we treat geographic, religious, and legal origins dummies as vectors of variables that enter the estimations together (e.g., all religious dummies enter the regression together instead of individually). Legal origins: German and Cont: North America are used as a base. 

(β) = the average β coefficient across all 8,129 estimations.

R Ratio = Robustness Ratio. If higher than 2, it suggests robustness (Young and Holsteen, 2015).

+ = % of models in which the variable enters with a positive sign.

+ & sig = % of models in which the variable enters with a positive & significant sign.

- = % of models in which the variable enters with a negative sign.

- & sig = % of models in which the variable enters with a negative & significant sign.

sign stab = sign stability indicating the percentage of models that have the same sign.

sig rate = significance rate indicating the percentage of models that report statistically significant coefficient. A significance rate of 95% or higher indicates "strong" robustness while a significance rate of 50% sets a lower bound for “weak” robustness (Raftery, 1995).

N = number of observations