

# What does “entrepreneurship” data *really* show?

Zoltan J. Acs · Sameeksha Desai · Leora F. Klapper

Accepted: 18 July 2008 / Published online: 9 September 2008  
© The Author(s) 2008. This article is published with open access at Springerlink.com

**Abstract** In this paper, we compare two datasets designed to measure entrepreneurship: The Global Entrepreneurship Monitor (GEM) dataset and the World Bank Group Entrepreneurship Survey (WBGES) dataset. We find a number of important differences when the data are compared. First, GEM data tend to report significantly higher levels of early stage entrepreneurship in developing economies than do the World Bank business entry data, while the World Bank business entry data tend to be higher than GEM data for developed countries. Second, we find that the magnitude of the difference between the datasets across countries is related to the local institutional and environmental conditions for entrepreneurs, after controlling for levels of economic development. Our findings suggest that entrepreneurs in developed countries have greater ease and incentives to incorporate, both for the benefits of greater access to formal financing and labor contracts, as well

as for tax and other purposes not directly related to business activities.

**Keywords** Entrepreneurship data · Global entrepreneurship monitor · World Bank · Entry · Institutions

**JEL Classifications** L26 · O10 · POO

## 1 Introduction

Since the mid-1980s, entrepreneurship has been increasingly considered as an important tool for economic growth and innovation across economies, regardless of stage of economic development. Entrepreneurship is now at the center of many policy questions related to science and technology, sustainability, poverty, human capital, endogenous resources, employment, regional and comparative advantages, etc. The surge of policy interest in entrepreneurship has, not surprisingly, been accompanied by growing academic research into its dynamics and processes. With respect to policy, research priorities have focused first on understanding (measuring) and second on creating environments supportive of entrepreneurship (Acs and Szerb 2007). One particularly important public policy issue for international development is the role played by institutional features of investment climate, for instance, the role of environmental conditions (Levie

---

Z. J. Acs (✉) · S. Desai  
School of Public Policy, George Mason University,  
Fairfax, VA, USA  
e-mail: zacs@gmu.edu

Z. J. Acs · S. Desai  
Max Planck Institute of Economics, Jena, Germany  
e-mail: desai@econ.mpg.de

L. F. Klapper  
Development Research Group, The World Bank,  
Washington, DC, USA  
e-mail: lklapper@worldbank.org

and Autio 2008) and indicators of the business environment (such as those measured by *Doing Business* reports—see World Bank 2007). For example, this includes measures of the regulatory burden for starting, operating and closing a business, such as the cost, number of days and number of procedures required to start a business.

In recent years, different sources of data on “entrepreneurship” have led to contradictory or inconclusive empirical findings for research into its dynamics.<sup>1</sup> For example, it is still unclear if—and in what direction—a causation exists between entrepreneurship and unemployment, poverty, taxation, regulatory burden, etc. Country-specific differences may certainly lead to contradictory findings, as well as the variety in the types of data used as broad measures of “entrepreneurship.” This has contributed to a great deal of confusion in entrepreneurship research. For this reason, it is critically important to understand what the data indicate, and exactly what element of entrepreneurial dynamics is being measured. The WBGES data, for example, measures the registration of LLCs, which is one kind of legal arrangement for a new firm. We discuss the implications of the various definitions of start-ups further in the comparative analysis section of our paper.

Separate studies using GEM and WBGES data have found contradictory results: While no relationship is found between GEM data and administrative barriers to starting a business, a significantly negative effect is found with WBGES data (van Stel et al. 2007 and Klapper et al. 2007, respectively).<sup>2</sup> It is possible that this—and similar contradictory results in the empirical entrepreneurship research—can be attributed to some degree to the differences in *what* the data capture. For this reason, we compare the two popular datasets designed to capture entrepreneurial dynamics.

In this paper, we compare the GEM dataset for early stage entrepreneurial activity and the WBGES dataset for formal business registration. We find two important trends when the data are compared descriptively. First, GEM data tend to report significantly higher levels of early stage entrepreneurship in developing economies than do the World Bank business entry data. Second, the World Bank business entry data tend to be higher than GEM data for developed countries.

There are at least three possible ways to interpret this discrepancy. First, the datasets simply measure different dynamics related to “entrepreneurship.” The WBGES measure rates of entry in the formal economy, and even more specifically, entry in the form of LLC establishments. The GEM data are perhaps more reflective of entrepreneurial intent and what some might call “entrepreneurial spirit.” For this reason, GEM data capture *informality* of entrepreneurship, particularly in developing countries. In particular, firm *formation* does not necessarily mean firm *registration*. Second, this discrepancy can also be interpreted as the spread between individuals who could *potentially* operate businesses in the formal sector—and those that *actually* do so. If this is the case, then GEM data may represent the potential supply of entrepreneurs, whereas WBGES data would represent the actual rate of entrepreneurship. This is interesting especially in the context of the allocation of talent (Murphy et al. 1991) and the allocation of entrepreneurship (Baumol 1990). In the allocation of talent model, the stock of talent is relatively constant, but its allocation towards a range of activities can change. Similarly, in the allocation of the entrepreneurship model, the stock of entrepreneurs in the economy is relatively constant, but the nature of their activities changes.

The motivation for entrepreneurs to operate in the formal versus informal sector is examined further in our empirical analysis. We find that the magnitude of differences reported in the datasets across countries is related to the institutional and environmental conditions for entrepreneurs. In terms of institutional differences, we find that the conditions related to registration, operation and closure of business are important; in terms of environmental differences, we find significant affects of economic and political conditions. Overall, entrepreneurs in developed countries have greater ease and incentives to incorporate, both for the benefits of greater access to formal financing and labor contracts,

<sup>1</sup> For a recent review of the literature, see National Research Council (2007) and Hoffmann and Oxholin (2006).

<sup>2</sup> This is also consistent with Klapper et al. (2006), who find a significant relationship between business registration in 35 European countries and entry barriers. De Soto (1990) and Djankov et al. (2002) find that costly regulations impede the setting up of businesses and stand in the way of economic growth. Djankov et al. (2002) find that high costs of entry exist in most countries, and that countries with more corruption have larger unofficial economies.

as well as for tax and other purposes not related to business activities. We elaborate on this further in the comparative analysis section of this paper.

## 2 Data description

### 2.1 GEM

The Global Entrepreneurship Monitor (GEM) project is unique in that while all countries collect official data on self-employment, the size distribution of firms, census data on all or most plants and firms, firm and plant entry, almost none of these registry sources are comparable across countries, even in developed countries. Official data sources differ in the way they define when an establishment enters a file and when it leaves, and how they handle self-employment makes cross-national comparisons almost impossible.<sup>3</sup> Therefore, one of the major strengths of the project is the application of uniform definitions and data collection across countries for international comparisons.

The intent of GEM is to systematically assess two things: the level of start-up activity or the prevalence of nascent firms and the prevalence of new or young firms that have survived the start-up phase. First, start-up activity (the “nascent” rate) is measured by the proportion of the adult population (18–64 years of age) in each country that is currently engaged in the process of creating a nascent business. Second, the proportion of adults in each country who are involved in operating a business that is less than 42 months old measures the presence of new firms (the “baby” rate). The distinction between nascent and new firms is made in order to determine the relationship of each to national economic growth. For both measures, the research focus is on entrepreneurial activity in which the individuals involved have a direct, but not necessarily full, ownership interest in the business.

### 2.2 World Bank group data

The goal of the *2007 World Bank Group Entrepreneurship Survey* was to collect a benchmark of formal entrepreneurial activity for a large number of developed and developing countries. The intent is that these

data will be used to compare private sector development across countries, as well as to monitor and evaluate the impact of regulatory reforms over time. In order to measure entrepreneurship and make data universally comparable, we developed a methodology that can be applicable across heterogeneous legal regimes and economic systems. Previous efforts had been made in this regard, but the great majority focused solely on the developed world and did not take into account differences in legal systems, sectors and economic structures (see United Nations 2005).

The WBGES defines the unit of measurement of entrepreneurship as:

Any economic unit of the formal sector incorporated as a legal entity and registered in a public registry, which is capable, in its own right, of incurring liabilities and of engaging in economic activities and transactions with other entities.

Notably, this definition excludes informal sector initiatives. This exclusion is based on the difficulties of quantifying the number of firms in the informal sector, rather than on its relevance for developing economies (Nielsen and Plovig 1997). The only way to measure the informal sector is through economic censuses, which due to their high costs are infrequently collected.

Furthermore, entrepreneurship is defined as:

The activities of an individual or a group aimed at initiating economic activities in the formal sector under a legal form of business.

However, few countries (i.e., Denmark) maintain “active” registries that annually confirm that registered firms are still operating. Therefore, official registration data include both businesses incorporated for economic activities, as well as those incorporated for tax or other non-business purposes (e.g., shell companies). An additional limitation of the data is that they do not report the number of closed businesses. The reasons differ from country to country, but are mainly due to the fact that the registrars generally have no enforcement mechanisms to oblige businesses to report closures. Although the number of closed companies is essential to paint a clear picture of the economic and entrepreneurial activities of a country, it is not yet feasible to obtain comparable data (Nucci 1999).

The WBGES database includes data on formal business registrations in 84 countries. The information was collected from business registries and other

<sup>3</sup> For a discussion of the GEM data, see Reynolds et al. (2005).

government sources via a survey and follow-up phone calls.<sup>4</sup> These other sources include statistical agencies, tax and labor agencies, chambers of commerce and private vendors (such as D&B), which were used only when business registry data were unavailable or non-existent.<sup>5</sup> The survey collected data on the year-end stock of total registered firms and new firms registered in the calendar year from 2003 to 2005.<sup>6</sup> Importantly, the definition of entrepreneurship includes only businesses that operate in the formal sector, and to maximize comparability across countries of different legal and economic systems, the database includes only limited liability corporations (LLCs).

For the purpose of the analysis in this paper, the data are used to calculate the “corporate” entrepreneurship rate, which is defined as the number of newly registered companies as a percentage of the adult population.

### 2.3 Comparative analysis

To compare entrepreneurship rates between the two databases, we calculate the spread between the “nascent” and “baby” entrepreneurship rates in GEM and the “corporate” entrepreneurship rate in WBGES.<sup>7</sup> The first new indicator, SPR\_N\_C,

measures the difference between percentages of individuals who in the process of starting a business (the GEM “nascent” rate) and those who have actually started a formal corporation. The second new indicator, SPR\_B\_C, measures the difference between the percentage of individuals operating a young business in either the formal or informal sector (“baby”), with the percentage of individuals who have chosen and/or succeeded in starting a formal corporation (“corporate”).

We interpret these spreads to reflect, in some part, a loss of potential formal sector participation. In other words, this can represent those individuals who were unsuccessful in registering their business because of barriers to registration that we later introduce or who chose to operate in the informal sector. The tendency of GEM data to be higher than WBGES data for developing countries is likely partly indicative of lost formal sector participation due to barriers to participation, and partly indicative of the informal economy due to choice. These are not mutually exclusive. In either case, the individual may still have started a business—but as we mentioned in the introduction, firm *formation* does not mean *registration*. We expect a higher spread—indicating a larger loss of entrepreneurial potential—in countries with weaker business environments.<sup>8</sup> The quality of the business environment, as measured by the Doing Business and other indicators, is collectively accepted as a critical determinant of entrepreneurial activity. These spreads, by country, are shown in Fig. 1.

What would we expect the data to show from a theoretical perspective? If the nascent rate represents early stage activity, we expect this to be higher than the young entrepreneurship rate. This is because many people that take “some steps” towards starting a business do not actually succeed. We also expect the young entrepreneurship rate to be larger than the formal rate, since many firms first are initially established under sole proprietorship, but incorporated at a later stage. In fact, for the United States, these rates are 8.12%, 4.98% and 2.55%, respectively. This does not, however, hold across developed and developing countries.

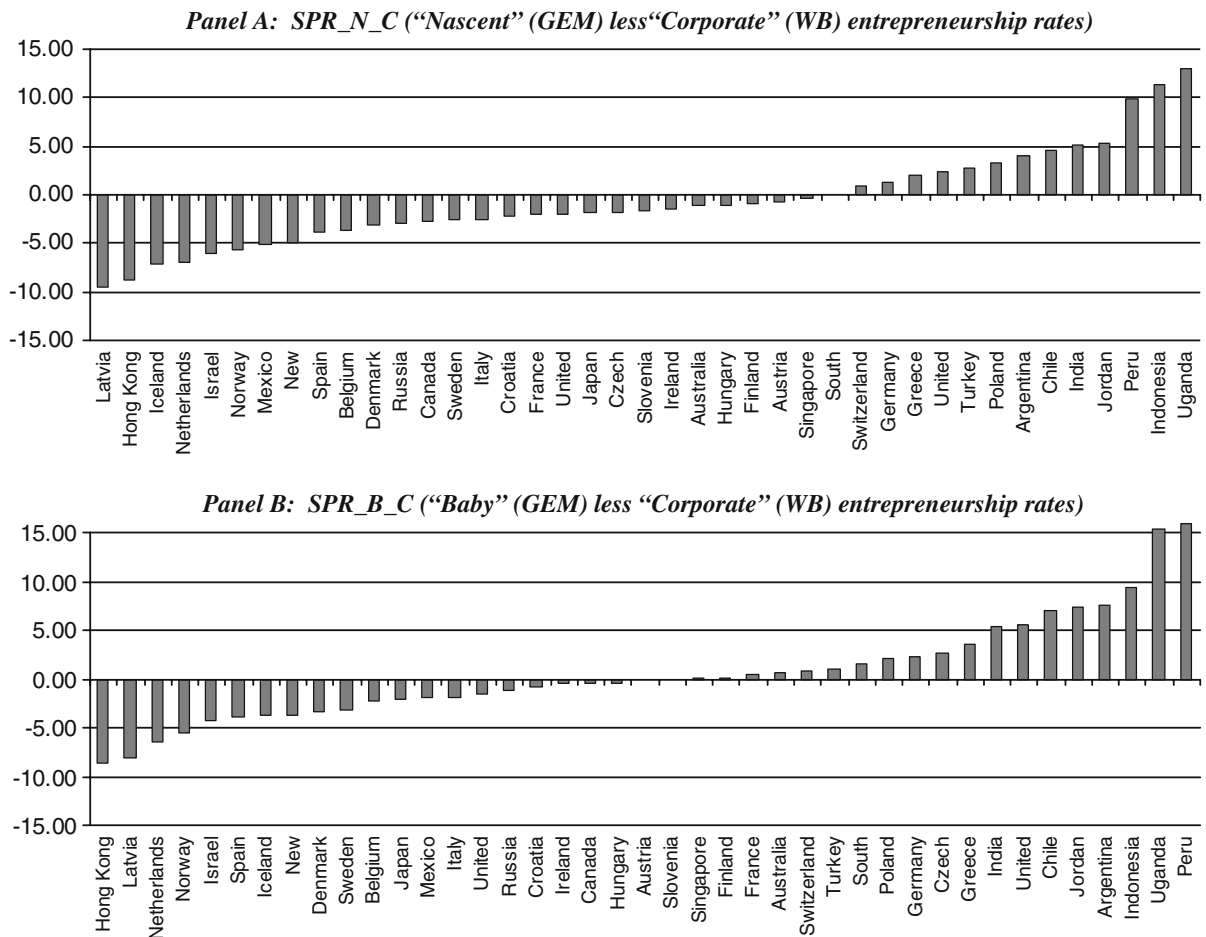
<sup>4</sup> The complete survey data and companion papers are available at: <http://www.ifc.org/ifcext/sme.nsf/Content/Entrepreneurship+Database>.

<sup>5</sup> A complete list of sources is provided in Klapper et al. (2007).

<sup>6</sup> For additional information, see Klapper and Quesada Delgado (2007) and Klapper et al. (2007).

<sup>7</sup> The SPR\_B\_C cannot be strictly compared. The nascent prevalence rate is for one point in time, so it is more or less an annual rate. However, the baby business data are for 42 months of activity, so it is not actually an ‘entry rate’ of new firms. The GEM data can be estimated for an annual rate. First, you need to estimate how many new births the numbers represent. Since there is an annual attrition rate at the end of 6 months, 95% of the firms would still be operating. For example, if 100 are born, this assumes that 95 will be operating at month 7. This increases the total count by 16% to compensate for the discontinuances. Second, we adjust from 42 months to 1 year. The final correction factor is 0.33. Using data from the 2004 US GEM Survey and using the US population base between 18 and 74, the mean baby business rate is 0.73 with a 95% confidence interval from 0.66 to 0.77. If you assume half are self-employed, as suggested by Census research, this is an annual birth rate of employer firms of about 0.36, with a 95% confidence interval from 0.33 to 0.38. We can compare this with the official US Census data for the US of about 0.36. So the GEM estimates are clearly within the limits of official US statistics ([http://www.sba.gov/advo/research/dyn\\_b\\_d8904.pdf](http://www.sba.gov/advo/research/dyn_b_d8904.pdf)).

<sup>8</sup> Data, by country, are shown in Annex A.



**Fig. 1** Nascent, young and formal entrepreneurship. Variables are defined in Table 1. *Panel A:*  $SPR\_N\_C$  [“nascent” (GEM) less “Corporate” (WB) entrepreneurship rates]. *Panel B:*  $SPR\_B\_C$  [“Baby” (GEM) less “Corporate” (WB) entrepreneurship rates]

In fact, it appears that in many countries—developed and developing—the young entrepreneurship rate and the nascent entrepreneurship rate are *less* than the formal entrepreneurship rate. This is the case not only in Hong Kong, but also in Latvia, The Netherlands, Norway, Israel, Iceland, New Zealand, Denmark, Sweden, Belgium, etc. In many developed countries, therefore, the formal entrepreneurship rate actually exceeds the young entrepreneurship rate and even the nascent entrepreneurship rate. Even within developing countries, the structure and types of entrepreneurial activities can differ (see Acs and Amorós 2008).

There are several possible explanations. In developing countries, a lower corporate rate might

actually represent a shift towards increased formalization of the economy. Newly registered companies may represent some aspect of formalization, where businesses that were not previously LLCs have newly converted their legal status. It is also important to note the unit of analysis is different in the datasets: GEM measures the number of individual entrepreneurs, possibly overlooking individuals that are involved in multiple new businesses. The WBGES dataset instead measures the number of businesses and can capture this dynamic. However, a possible complication also results from the WBGES measure: Formal entrepreneurship includes both actual businesses *and* LLCs that are a legal vehicle for purposes other than

starting a new business. For instance, entrepreneurs might use registrations to achieve other business ends such as reducing taxes (e.g., shell companies) and avoiding regulatory burdens (e.g., labor laws).<sup>9</sup> For example, in the United States, firms may register several LLCs as a way to limit liability for different lines of businesses. In Hong Kong, where the formal rate far surpasses the young business formation rate, all real estate sales are first converted to an LLC to avoid taxes. The incentive to register firms for redundant or non-business activities might be greater in developed countries with more complex (and enforced) tax and regulatory structures.

## 2.4 Data and summary statistics

The sample for the analysis is a pooled, cross-sectional, longitudinal unbalanced panel of 90 observations across 40 countries with non-missing explanatory variables in both the GEM and WBES databases for 2003, 2004 and 2005.<sup>10</sup> Summary statistics are shown in Table 1. The mean spread with nascent entrepreneurs (SPR\_N\_C) is  $-0.36\%$ , and the spread with young firms (SPR\_B\_C) is  $-1.55\%$ , which suggests that on average the two measures are very similar. However, we find a standard deviation of over 4% for both indicators—maximum values of over 9% and minimum values less than  $-9\%$ —and variation across economic and political environments.

We consider a variety of country characteristics as predictors of entrepreneurial activity, which vary over time. We include log GDP per capita (GDPPC) in all estimations to control for economic development because of the varied levels of development of countries for which we have data. As an additional explanatory variable, we include the ratio of domestic credit to the private sector as a percentage of GDP as a measure of financial development (DomCredit).

We use four measures of the regulatory barriers: first, an indicator of the difficulty of hiring and firing employees (Labor\_Rig); second, the log cost of business registration (Entry\_Cost); third, the log number of procedures required to start a business (Entry\_Proc); fourth, the ease of closing a business, proxied by the estimated recovery rate claimants can expect following foreclosure or bankruptcy (Rec\_Rate). These measures indicate the difficulties in starting, operating and closing a business.

It is important to note that these indicators measure the barriers for a “typical” *formal sector* firm, which might in part explain the weak relationship with GEM data. For instance, the methodology for entry barriers assumes:

“The business is:

- A limited liability company.
- Has start-up capital of ten times income per capita at the end of 2005, paid in cash.
- Has a turnover of at least 100 times income per capita.”<sup>11</sup>

We expect that these barriers would have a stronger relationship with the formal entrepreneurship rates in the WB database. Furthermore, these indicators might be important predictors of a firm’s decision to operate in the formal versus informal sector.

Next, we include indicators of operational risk, which may proxy for the risks and benefits of individuals of operating a firm in the formal (rather than informal) sector. For instance, we would expect individuals to be less willing to operate illegally (and more likely to pay taxes) in countries where registration laws are enforced, corruption is lower, and the economy is healthy. First, we include an index of political risk (Pol\_Risk), which measures corruption, government stability, etc. Second, we include an index of law and order (Law\_Order), which measures the efficiency of the legal and judicial system. Third, we include an index of economic risk (Econ\_Risk), which measures the economic growth of the country. Fourth, we include a composite risk index, which is

<sup>9</sup> For instance, laws on hiring and firing employees in Italy applies only to firms with more than 15 employees, which might encourage business owners to register multiple smaller firms (Klapper et al. 2006).

<sup>10</sup> The complete list of countries is shown in Annex B.

<sup>11</sup> <http://www.doingbusiness.org/MethodologySurveys/StartingBusiness.aspx>.



**Table 1** Variable definitions and summary statistics

Variable	Obs.	Description	Mean	SD
SPR_N_C	90	The spread between the “nascent” entrepreneurship rate (GEM)—defined as the number of people actively involved in starting a new venture, as a percentage of adult population—and “corporate” entrepreneurship—defined as the percentage of newly registered limited-liability firms (<1 year), as a percentage of adult population. A higher value indicates a greater loss of entrepreneurial potential	−0.36	4.14
SPR_B_C	90	The spread between the “baby” entrepreneurship rate (GEM)—defined as the number of people that are owners/managers of a business that is <42 months old, as a percentage of adult population—and “corporate” entrepreneurship. A higher value indicates a greater loss of entrepreneurial potential	−1.55	3.69
GDPPC	90	Log GDP per capita (WBI statistics)	9.86	0.65
DomCredit	90	Domestic credit to the private sector as a percentage of GDP (WBI statistics)	98.31	55.63
Labor_Rig	88	The “Rigidity of Employment” index, calculated as the average of three subindices: a difficulty of hiring index, a rigidity of hours index and a difficulty of firing index; normalized between 0 and 1. A higher value indicates greater employment barriers (DB)	0.34	0.19
Entry_Cost	88	Log cost of business registration, expressed as a percentage of per capita GNP. A higher value indicates greater entry barriers (DB)	1.96	1.17
Entry_Proc	88	Log number of procedures to register a business. A higher value indicates greater entry barriers (DB)	2.01	0.48
Rec_Rate	88	The log estimate of how many cents on the dollar claimants—creditors, tax authorities and employees—recover from an insolvent firm, as a measure of the efficiency of foreclosure or bankruptcy procedures. A higher value indicates lower closure barriers (DB)	0.46	0.17
Pol_Risk	90	An index of political risk, measured as the average of 12 subindices, including government stability, internal and external conflict, corruption, law and order, and bureaucracy quality; normalized between 0 and 1. A higher value indicates lower risk (ICRG)	0.81	0.09
Law_Order	90	An index of the strength and impartiality of the legal system and popular observance of the law; normalized between 0 and 1. A higher value indicates lower risk (ICRG)	0.82	0.18
Econ_Risk	90	An index of economic risk, indicating a country’s current economic strengths and weaknesses, measured as the average of five subindices: GDP per capita, real GDP growth, inflation, budget balance as a percentage of GDP, and current account as a percentage of GDP; normalized between 0 and 1. A higher value indicates lower risk (ICRG)	0.80	0.07
Comp_Risk	90	A composite index of political, economic and financial risk ratings (where Financial_Risk is a country’s ability to finance its official, commercial and trade debt obligations); normalized between 0 and 1. A higher value indicates lower risk (ICRG)	0.80	0.07
Informal	40	Share of the informal economy, calculated as the size of the informal economy as a percentage of official GNI; normalized between 0 and 1 (DB)	0.23	0.11

The sample is a pooled, cross-sectional, longitudinal unbalanced panel across 41 countries with non-missing explanatory variables for 2003, 2004 and 2005. “GEM” is the Global Entrepreneurship Monitor; “WBGED” is the World Bank Group Entrepreneurship Database; “DB” is the World Bank Doing Business Database ([www.doingbusiness.org](http://www.doingbusiness.org)); “ICRG” is the International Country Risk Guide

an average of political, economic and governmental financial risk and stability.

### 3 Empirical results

Table 2 shows the correlation matrix of our variables. Univariate tests show significance with all variables except employment laws. An explanation might be that both formal and informal young firms are less

likely to hire a large number of employees.<sup>12</sup> Because of the large and significant correlation between the explanatory variables, estimations are run separately, while controlling for economic development through logGDP per capita.

<sup>12</sup> This is especially true since formal firms in developing countries are likely to be in the sectors of wholesale and retail trade—and unlikely to be in manufacturing—which are less dependent on labor (Klapper et al. 2007).

**Table 2** Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
SPR_N_C (1)	1.00										
SPR_B_C (2)	0.93	1.00									
GDPPC (3)	−0.52	−0.53	1.00								
DomCredit (4)	−0.32	−0.35	0.60	1.00							
Labor_Rig (5)	0.00	0.03	−0.09	−0.39	1.00						
Entry_Cost (6)	0.39	0.41	−0.52	−0.39	0.48	1.00					
Entry_Proc (7)	0.44	0.46	−0.50	−0.36	0.48	0.75	1.00				
Rec_Rate (8)	−0.55	−0.54	0.61	0.48	−0.39	−0.56	−0.60	1.00			
Pol_Risk (9)	−0.52	−0.50	0.73	0.48	−0.26	−0.62	−0.69	0.76	1.00		
Law_Order (10)	−0.51	−0.50	0.53	0.40	−0.32	−0.60	−0.69	0.67	0.81	1.00	
Econ_Risk (11)	−0.44	−0.43	0.59	0.28	−0.31	−0.60	−0.58	0.66	0.63	0.59	1.00
Comp_Risk (12)	−0.58	−0.53	0.69	0.43	−0.28	−0.57	−0.62	0.77	0.88	0.74	0.83

Variables are defined in Table 1. All coefficients—with the exception of Labor\_Rig—are significant at 1%

Figure 2 shows scatter plots and univariate tests of our explanatory variables. We find significant relationships for both the SPR\_N\_C and SPR\_B\_C. As expected, the spread between the two measures is negatively related to per capita GDP, composite risk, recovery rate and law and order. It is positively related to the number of procedures needed to register a business and the share of the informal economy.

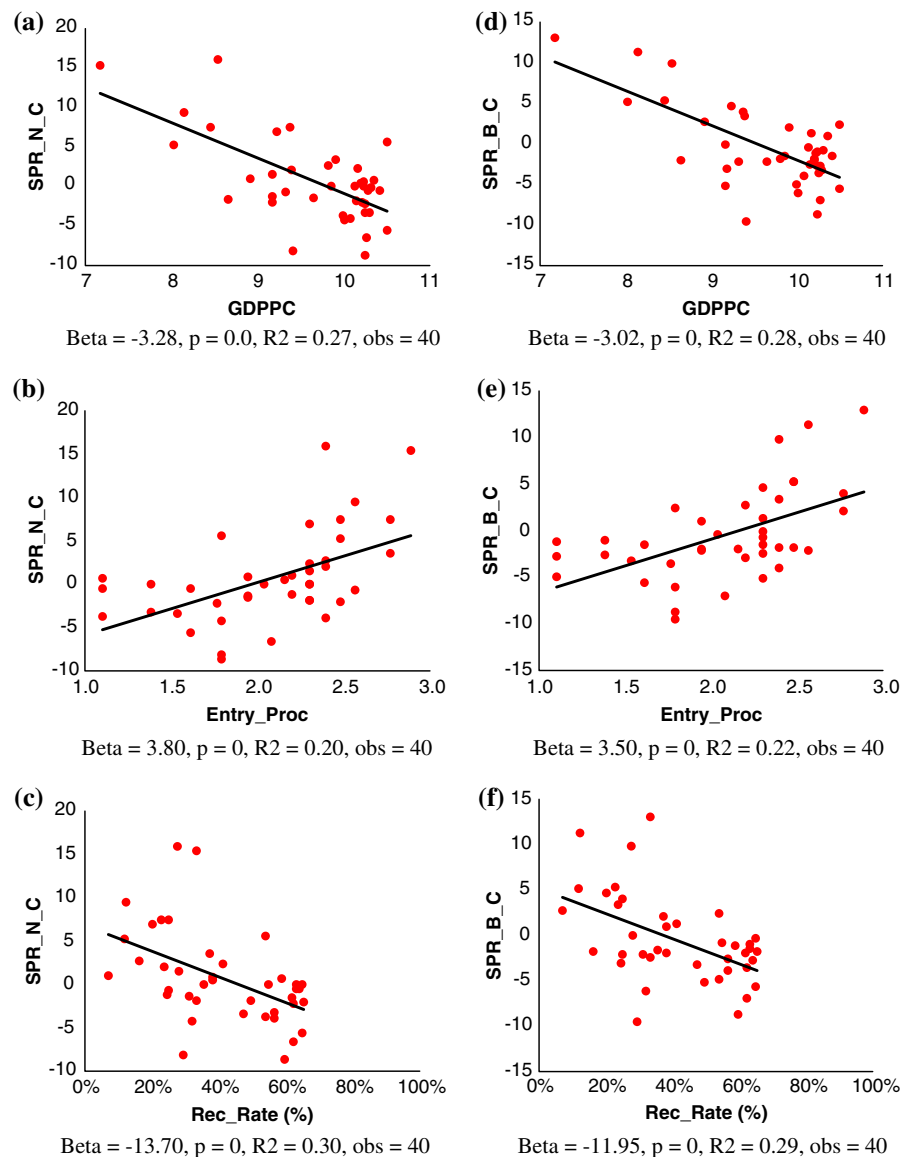
Table 3 shows our estimation results for the spread between nascent and formal entrepreneurship. We find no relationship between this spread and domestic credit, which might suggest that start-ups are less dependent on formal bank financing (and depend more on personal savings). The strongest relationship among our investment climate variables is with closure costs—since the default rate of new firms is very high, firms that expect to get the lowest return on their investment might be least likely to undertake the time and cost of joining the formal sector (and benefiting from formal legal bankruptcy proceedings). We find the interaction terms of entry costs, entry procedures and recovery rates with GDP per capita to be significant barriers to starting (and closing) a business matter more in lower-income countries. Or, in other words, individuals in developing countries are only likely to have incentives to join the formal sector if entry barriers are low. A possible explanation is that many developing countries host

substantial informal sectors, so entrepreneurs are able to operate entirely within the informal economy. For example, the ILO estimates 60 percent of the workforce in Asia to be in the informal sector (ILO 2007). Individuals can start businesses that meet demand, and derive supply, within the informal sector. In such cases, they have little actual need to join the formal sector in order to operate.

Table 4 shows the relationship between the spread with nascent entrepreneurs and measures of country risk. We find a strong and significant relationship with the composite risk index—again, individuals are more likely to choose and succeed in joining the formal sector if the political, economic and financial risks are low. Furthermore, the interaction with law and order is significant.

Next, we use as our dependent variable the spread between young business—both formal and informal—and formal entrepreneurship. We expect this spread to be the largest in countries with weaker business environments (and larger informal sectors). Table 5 shows that in this case, in addition to recovery rates, entry procedures (and the interaction with GDP per capita) is significant, i.e., entry barriers matter. Table 6 shows that law and order—legal and judicial efficiency—is the most important determinant in the decision whether or not to operate in the formal sector and/or to register as a limited-liability company.



**Fig. 2** Scatter plots of “Potential” entrepreneurship

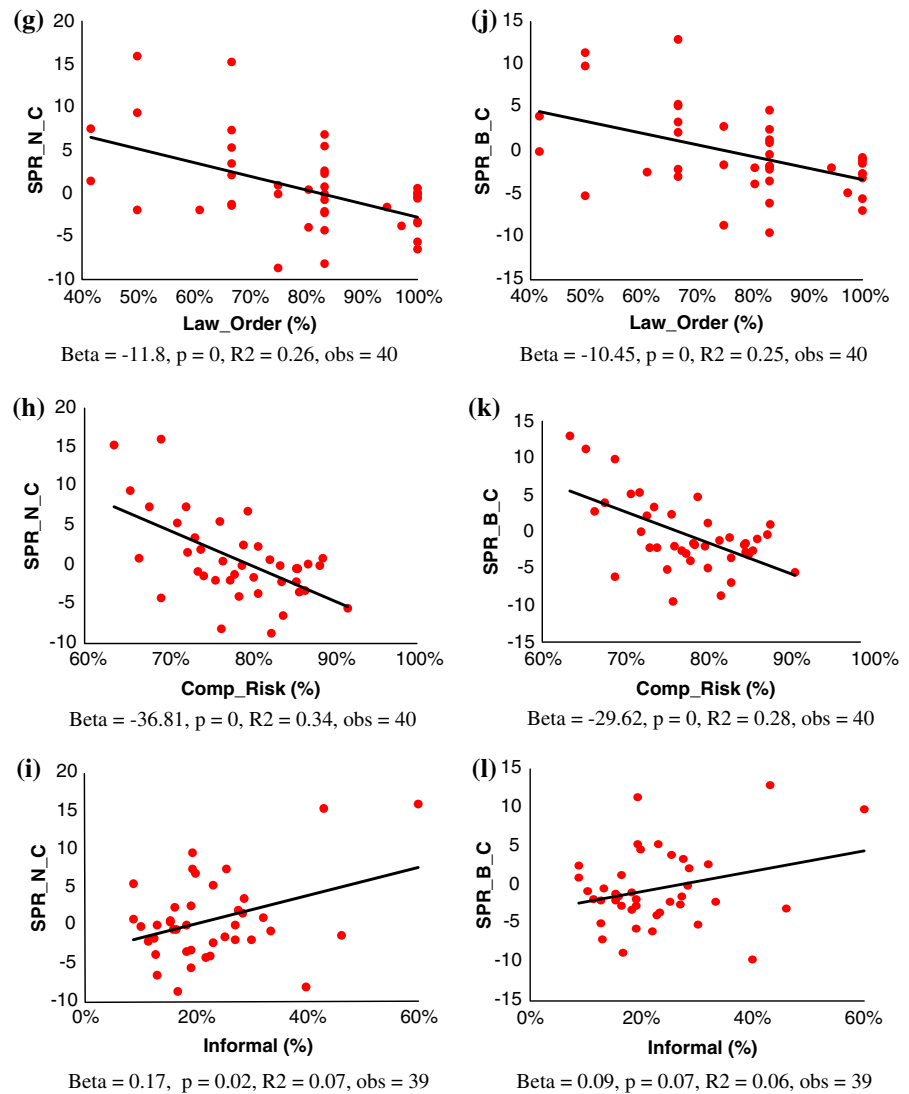
The results raise one interesting question. As entry barriers increase, the spread between the informal and the formal sector rises, as expected, and as entry procedures fall, the spread between the formal and informal sector falls. The implication is that barriers to entry are greater for corporate entrepreneurship than for young businesses that have not incorporated or for nascent entrepreneurs where they are in the process of starting a business. However, in developed countries, the spread between the informal and formal sectors not only decreases, but is often positive; i.e., the number of limited-liability companies is greater

than the sum of sole proprietors and informal firms. This implies that it is at least as easy to start a limited liability company as a sole proprietorship.

#### 4 Conclusion

The purpose of this paper is to compare two datasets designed to capture entrepreneurial dynamics: the GEM data for early stage entrepreneurial activity and the World Bank Entrepreneurship Group dataset for formal business registration. We find a number of

Fig. 2 continued



important differences in the data. First, the GEM data tend to report significantly lower levels of early stage entrepreneurial activity in developed countries. In other words, it is more common to start a formal business in a developed country than a sole proprietorship. Second, the GEM data tend to be higher for developing countries than for developed countries. One possible explanation is the distinction between intent and informality of entrepreneurial

activity particularly in developing countries that is captured by GEM data. However, important exceptions to this are found for both the United States and Germany in particular. This suggests that firms in developed countries have greater ease and incentives to incorporate, both for the benefits of greater access to formal financing and labor contracts, as well as for tax and other purposes not related to business activities.

**Table 3** The effect of the investment climate on “potential” nascent entrepreneurship

	DomCredit	DomCredit	Labor_Rig	Labor_Rig	Entry_Cost	Entry_Cost	Entry_Proc	Entry_Proc	Rec_Rate	Rec_Rate
GDPPC	-3.77 (0.00)***	-4.12 (0.00)***	-3.72 (0.00)***	-2.75 (0.03)**	-3.16 (0.00)***	0.35 (0.84)	-3.13 (0.00)***	8.07 (0.06)*	-2.56 (0.00)***	-5.16 (0.00)***
DomCredit	0.00 (0.84)	-0.12 (0.49)								
DomCredit*GDPPC		0.01 (0.48)								
Labor_Rig			-1.00 (0.74)	31.62 (0.38)						
Labor_Rig*GDPPC				-3.35 (0.36)						
Entry_Cost					0.51 (0.37)	10.88 (0.02)**				
Entry_Cost*GDPPC						-1.07 (0.02)**				
Entry_Proc							1.67 (0.18)	47.29 (0.01)***		
Entry_Proc*GDPPC								-4.50 (0.01)***		
Rec_Rate									-7.41 (0.04)**	-78.69 (0.06)*
Rec_Rate*GDPPC										7.35 (0.09)*
Year	-0.19 (0.33)	-0.20 (0.32)	-0.17 (0.36)	-0.17 (0.40)	-0.18 (0.33)	-0.24 (0.22)	-0.17 (0.38)	-0.24 (0.24)	-0.21 (0.33)	-0.25 (0.25)
Constant	423.05 (0.28)	448.64 (0.27)	386.67 (0.32)	372.14 (0.37)	389.83 (0.29)	474.76 (0.22)	363.67 (0.34)	395.76 (0.33)	450.77 (0.30)	547.71 (0.21)
Observations	90	90	88	88	88	88	88	88	88	88

Variables are defined in Table 1. The sample is a pooled, cross-sectional, longitudinal unbalanced panel across 41 countries with non-missing explanatory variables for 2003, 2004 and 2005. The dependent variable is “SPR\_B\_C,” which is a measure of lost entrepreneurial potential. The regressions are estimated using population-averaged generalized estimating equations (GEE), with a year trend added as a control. Z-scores are shown in brackets beneath regression coefficients. Asterisks, \*, \*\* and \*\*\*, indicate significance at 10%, 5% and 1%, respectively

**Table 4** The effect of country risks on “potential” nascent entrepreneurship

	Pol_Risk	Pol_Risk	Law_Order	Econ_Risk	Comp_Risk	Comp_Risk
GDPPC	-2.66 (0.01)***	-9.73 (0.02)**	-3.20 (0.00)***	-12.36 (0.00)***	-3.1036 (0.00)***	-6.6565 (0.42)
Pol_Risk	-10.22 (0.18)	-109.17 (0.06)*				-2.27 (0.01)***
Pol_Risk*GDPPC		10.16 (0.09)*				-11.10 (0.07)*
Law_Order			-3.40 (0.24)	-119.67 (0.00)***		
Law_Order*GDPPC				12.17 (0.00)***		
Econ_Risk				-4.6381 (0.48)		-50.959 (0.65)
Econ_Risk*GDPPC						11.8484 (0.25)
Comp_Risk					-20.42 (0.02)**	-141.51 (0.09)*
Comp_Risk*GDPPC						12.26 (0.14)
Year	-0.25 (0.27)	-0.28 (0.21)	-0.16 (0.45)	-0.24 (0.16)	-0.2058 (0.35)	-0.30 (0.21)
Constant	537.36 (0.24)	655.96 (0.14)	359.62 (0.41)	604.85 (0.09)*	445.2772 (0.31)	358.0772 (0.42)
Observations	90	90	90	90	90	90

Variables are defined in Table 1. The sample is a pooled, cross-sectional, longitudinal unbalanced panel across 41 countries with non-missing explanatory variables for 2003, 2004 and 2005. The dependent variable is “SPR\_B\_C,” which is a measure of lost entrepreneurial potential. The regressions are estimated using population-averaged generalized estimating equations (GEE), with a year trend added as a control. Z-scores are shown in brackets beneath regression coefficients. Asterisks, \*, \*\* and \*\*\*, indicate significance at 10%, 5% and 1%, respectively

**Table 5** The effect of the investment climate on “potential” young entrepreneurship

	DomCredit	DomCredit	Labor_Rig	Labor_Rig	Entry_Cost	Entry_Cost	Entry_Proc	Entry_Proc	Rec_Rate	Rec_Rate
GDPPC	-3.15 (0.00)***	-3.36 (0.00)***	-3.41 (0.00)***	-2.39 (0.03)**	-2.62 (0.00)***	0.20 (0.89)	-2.69 (0.00)***	6.25 (0.09)*	-2.38 (0.00)***	-5.23 (0.00)***
DomCredit	0.00 (0.57)	-0.07 (0.65)								
DomCredit*GDPPC		0.01 (0.67)								
Labor_Rig			-0.44 (0.86)	34.73 (0.27)						
Labor_Rig*GDPPC				-3.60 (0.26)						
Entry_Cost					0.75 (0.12)	9.28 (0.02)**				
Entry_Cost*GDPPC						-0.88 (0.03)**				
Entry_Proc							2.10 (0.06)*	38.73 (0.01)***		
Entry_Proc*GDPPC								-3.63 (0.01)**		
Rec_Rate									-6.77 (0.03)**	-83.99 (0.02)**
Rec_Rate*GDPPC										7.96 (0.03)**
Year	-0.22 (0.29)	-0.23 (0.28)	-0.22 (0.34)	-0.22 (0.35)	-0.22 (0.31)	-0.28 (0.21)	-0.21 (0.35)	-0.28 (0.23)	-0.26 (0.27)	-0.30 (0.19)
Constant	476.91 (0.26)	497.12 (0.25)	467.08 (0.30)	454.92 (0.33)	458.09 (0.29)	555.27 (0.21)	433.24 (0.33)	488.84 (0.29)	542.14 (0.24)	646.95 (0.15)
Observations	90									

Variables are defined in Table 1. The sample is a pooled, cross-sectional, longitudinal unbalanced panel across 41 countries with non-missing explanatory variables for 2003, 2004 and 2005. The dependent variable is “SPR\_B\_C,” which is a measure of lost entrepreneurial potential. The regressions are estimated using population-averaged generalized estimating equations (GEE), with a year trend added as a control. Z-scores are shown in brackets beneath regression coefficients. Asterisks, \*, \*\* and \*\*\*, indicate significance at 10%, 5% and 1%, respectively

**Table 6** The effect of country risks on “potential” young entrepreneurship

	Pol_Risk	Pol_Risk	Law_Order	Econ_Risk	Econ_Risk	Comp_Risk	Comp_Risk
GDPPC	-2.82 (0.00)***	-8.88 (0.02)**	-2.65 (0.00)***	-8.49 (0.00)***	-11.9016 (0.12)	-3.3218 (0.00)***	-2.49 (0.00)***
Pol_Risk	-5.70 (0.41)	-90.39 (0.08)*					
Pol_Risk*GDPPC		8.69 (0.10)*					
Law_Order			-5.20 (0.05)**	-81.34 (0.03)**			
Law_Order*GDPPC				7.88 (0.04)**			
Econ_Risk					-123.2368 (0.23)	-5.8403 (0.40)	
Econ_Risk*GDPPC						4.5009 (0.69)	
Comp_Risk						-13.11 (0.10)	-139.42 (0.07)*
Comp_Risk*GDPPC							12.78 (0.09)*
Year	-0.28 (0.21)	-0.31 (0.16)	-0.21 (0.36)	-0.26 (0.23)	-0.2223 (0.29)	-0.1365 (0.52)	-0.33 (0.14)
Constant	596.98 (0.18)	702.36 (0.11)	445.60 (0.33)	601.45 (0.16)	566.1512 (0.19)	310.607 (0.47)	669.41 (0.14)
Observations	90	90	90	90	90	90	90

Variables are defined in Table 1. The sample is a pooled, cross-sectional, longitudinal unbalanced panel across 41 countries with non-missing explanatory variables for 2003, 2004 and 2005. The dependent variable is “SPR\_B\_C,” which is a measure of lost entrepreneurial potential. The regressions are estimated using population-averaged generalized estimating equations (GEE), with a year trend added as a control. Z-scores are shown in brackets beneath regression coefficients. Asterisks, \*, \*\* and \*\*\*, indicate significance at 10%, 5% and 1%, respectively



**Acknowledgements** The opinions expressed do not necessarily represent the views of the World Bank, its Executive Directors or the countries they represent. We thank Paul Reynolds for valuable comments. We received valuable feedback from participants at the 2007 GEM Conference and 2008 Babson College Entrepreneurship Research Conference.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Annex A: Nascent, young and formal entrepreneurship

Shown are averages of non-missing variables for 2003, 2004 and 2005. “Nascent” is the number of people actively involved in starting a new venture, as a percentage of adult population; “baby” is the number of people that are owners/managers of a business that is <42 months old, as a percentage of adult population, and “corporate” is the percentage

Country	“Nascent”	“Young”	“Formal”	SPR_B_C	SPR_N_C
Argentina	9.17	5.65	1.67	3.98	7.50
Australia	7.32	5.58	6.70	−1.12	0.61
Austria	3.02	2.37	3.10	−0.73	−0.08
Belgium	2.64	1.25	4.83	−3.58	−2.19
Canada	5.88	3.66	6.35	−2.69	−0.47
Chile	8.49	6.23	1.58	4.65	6.91
Croatia	2.84	1.49	3.60	−2.11	−0.76
Czech Republic	6.41	1.98	3.77	−1.79	2.64
Denmark	2.68	2.86	6.04	−3.18	−3.36
Finland	3.29	2.26	3.24	−0.98	0.05
France	3.47	1.02	3.00	−1.98	0.47
Germany	3.16	2.31	0.84	1.27	2.34
Greece	3.92	2.54	0.43	2.10	3.49
Hong Kong	1.61	1.58	10.29	−8.71	−8.68
Hungary	2.96	2.28	3.35	−1.07	−0.40
Iceland	7.83	4.46	11.64	−7.18	−3.81
India	5.42	5.31	0.10	5.21	5.32
Indonesia	9.63	11.51	0.18	11.33	9.45
Ireland	5.05	4.03	5.56	−1.53	−0.51
Israel	4.32	2.53	8.59	−6.06	−4.27
Italy	2.49	1.90	4.37	−2.47	−1.87
Japan	0.96	1.21	3.02	−1.81	−2.06
Jordan	10.38	8.26	2.94	5.32	7.44
Latvia	4.17	2.77	12.33	−9.56	−8.16
Mexico	4.59	1.36	6.54	−5.18	−1.95
The Netherlands	2.43	2.01	8.96	−6.94	−6.53
New Zealand	9.02	7.82	12.73	−4.92	−3.71
Norway	4.14	4.11	9.69	−5.58	−5.55
Peru	31.36	12.93	3.05	9.88	16.00
Poland	3.92	5.20	1.85	3.35	2.07
Russia	3.46	1.71	4.69	−2.98	−1.23
Singapore	3.33	2.98	3.03	−0.39	0.02
Slovenia	2.62	1.08	2.64	−1.56	−0.02

**Annex A** continued

Country	“Nascent”	“Young”	“Formal”	SPR_B_C	SPR_N_C
South Africa	3.40	1.79	1.86	−0.07	1.54
Spain	2.95	2.97	6.90	−3.93	−3.95
Sweden	1.81	2.37	5.02	−2.64	−3.21
Switzerland	3.49	3.71	2.71	1.00	0.78
Turkey	2.20	4.01	1.25	2.76	0.95
Uganda	16.01	18.02	0.66	13.00	15.35
United Kingdom	3.41	3.07	5.01	−1.94	−1.60
United States	8.12	4.98	2.55	2.43	5.57

**Annex B: List of countries**

Argentina  
 Australia  
 Austria  
 Belgium  
 Canada  
 Chile  
 Croatia  
 Czech Republic  
 Denmark  
 Finland  
 France  
 Germany  
 Greece  
 Hong Kong  
 Hungary  
 Iceland  
 India  
 Indonesia  
 Ireland  
 Israel  
 Italy  
 Japan  
 Jordan  
 Latvia  
 Mexico  
 The Netherlands  
 New Zealand  
 Norway  
 Peru  
 Poland  
 Russia

**Annex B** continued

Singapore  
 Slovenia  
 South Africa  
 Spain  
 Sweden  
 Switzerland  
 Turkey  
 Uganda  
 United Kingdom  
 United States

of newly registered limited-liability firms (<1 year), as a percentage of adult population. SPR\_N\_C is the spread between nascent and formal entrepreneurship rates, and “SPR\_B\_C” is the spread between Young and Formal entrepreneurship rates.

**References**

- Acs, Z. J., & Amoros, J. E. (2008). Entrepreneurship and competitiveness dynamics in Latin America. *Small Business Economics*, 3(3), this issue. doi: [10.1007/s11187-008-9133-y](https://doi.org/10.1007/s11187-008-9133-y).
- Acs, Z. J., & Szerb, L. (2007). Entrepreneurship, economic growth and public policy. *Small Business Economics*, 28, 109–122.
- Baumol, W. J. (1990). Entrepreneurship: Productive, unproductive and destructive. *Journal of Political Economy*, 98(5), 893–921.
- De Soto, H. (1990). *The other path*. New York: Harper and Row.

- Djankov, S., La Porta, R., de-Silanes, F., & Shleifer, L. A. (2002). The regulation of entry. *Quarterly Journal of Economics*, 117, 1–35.
- Hoffmann, A., & Oxholm, A. S. Quality manual of entrepreneurship indicators-version 2 October 2006, FORA, Denmark, from <http://www.foranet.dk/Publikationer/Rapporter/Data/Quality%20Assessment%20of%20Entrepreneurship%20Indicators%20Version%202.aspx>.
- International Labour Organization. (2007). In Asia, informal work shifts but remains massive, from [www.ilo.org](http://www.ilo.org)
- Klapper, L., Amit, R., Mauro, F. G., & Delgado, M. J. Q. (2007). Entrepreneurship and firm formation across countries. Development Research Group Working Paper No. XX, World Bank, Washington D.C.
- Klapper, L., Laeven, L., & Rajan, R. (2006). Entry regulation as a barrier to entrepreneurship. *Journal of Financial Economics*, 82(3), 591–629.
- Klapper, L., & Quesada Delgado, J. M. (2007). Entrepreneurship: New data on business creation and how to promote it. Viewpoint Series, Note XX, World Bank Group Financial Sector Vice Presidency, Washington D.C.
- Levie, J., & Autio, E. (2008). A theoretical grounding and test of the GEM model. *Small Business Economics*, 31(3), this issue. doi:10.1007/s11187-008-9136-8.
- Murphy, K. M., Shleifer, A., & Vishny, R. W. (1991). The allocation of talent: Implications for growth. *The Quarterly Journal of Economics*, 106(2), 503–530.
- National Research Council. (2007). *Understanding business dynamics: An integrated data system for America's future*. Washington, D. C: The National Academies Press.
- Nielsen, P. B., & Plovig, J. (1997). Concepts used in statistical business registers in view of globalization and the informal economy. *Statistics Denmark, International Statistical Review*, 351–363.
- Nucci, A. (1999). The demography of business closings. *Small Business Economics*, 12(1), 25–39.
- PRS Group. (2007). International country risk guide. The PRS Group: New York.
- Reynolds, P., Bosma, N., Autio, E., Hunt, S., De Bono, N., Servais, I., et al. (2005). Global entrepreneurship monitor: Data collection design and implementation, 1998–2003. *Small Business Economics*, 24(3), 205–231.
- United Nations. (2005). Recommandations Internationales Concernant Les Statistiques de L'activité économique. E/ESCAP/CPR(2)9, Commission économique et Sociale pour L'Asie et le Pacifique, November.
- van Stel, A., Storey, D. J., & Thurik, A. R. (2007). The effect of business regulations on nascent and young business entrepreneurship. *Small Business Economics*, 28, 171–186.
- World Bank. (2007). *Doing business 2007—How to reform*. Washington D. C: The World Bank.