APPENDIX

A Variables of the model

Table A1 Description of variables

Variable	Description
Age	Describes the age range in which microentrepreneurs is (18 - 30), (31 - 42), (43 - 58) and (more than 58)
Education	Describes the range of the education level of the microentrepreneurs: primary and others, high school or secondary and university
Female	One if the microentrepreneurs is a woman and zero otherwise
Number of Workers	Besides you, how many people on average have worked at this establishment in the last 12 months or in the months of operation?
Internet use	One if firm uses internet or zero if not to carry out your economic activities related to your business
Financial inclusion index	Where zero corresponds to complete exclusion and one complete inclusion
Separate Account	One if there is an exclusive savings account for this business, zero otherwise
Use electronic wallet	One if the firm has ever used an electronic wallet, zero otherwise
Bank Loan	One if a bank loan or credit was requested and accepted with traditional banks (eg BBVA, Bancolombia, etc.), zero otherwise
Family Insurance	One if the firm currently has any type of insurance for your family, for you or for your business, zero otherwise
Formality index	Where zero corresponds to complete informality and one corresponds to complete formality
Operating Permit	One if the firm has an operating permit, zero otherwise
Accounting records	One if the firm has accounting records, through formal accounting or maintaining personal records, zero otherwise
Commercial Registry	One if it has a commercial register or zero otherwise
Tax registry	One if the firm is either register with the tax authority as a firm or as a natural person, zero otherwise
Insured Workers	Proportion of workers with social security benefits

Table A2 Variables of Business Practices index

Business Practices	One corresponds if the microbusiness adopts all business practices and zero if it does not adopt any of them
marketing	
Business practice 1	One if firm visited a competitor's business to learn their prices, zero otherwise
Business practice 2	One if firm visited a competitor's business to see the products offered, zero otherwise
Business practice 3	One if firm asked his current customers if there was any product or service that they would like to buy or have in his business, zero otherwise
Business practice 4	One if the business asked a vendor about what products are selling well in the industry or sector, zero otherwise
Business practice 5	One if firm attract customers with special offers, zero otherwise
Business practice 6	One if firm did some kind of marketing, zero otherwise
Business practice 7	One if firm suggested new products or services to his customers, zero otherwise
Business practice 8	One if firm has weekly sales goals or objectives or monthly, zero otherwise
Business practice 9	One if it uses the internet, books, magazines or newspapers about new trends in the sector, zero otherwise
Inventory	
Business practice 10	One if it tried to negotiate with a provider for lower prices for your raw materials, zero other
Business practice 11	One if it compared prices and / or quality offered with alternative suppliers or raw material sources, zero otherwise
Business practice 12	One if firm keeps inventories, zero otherwise
Business practice 13	One if firm keeps inventory records of final products and raw materials, zero otherwise
Sales and Purchases	
Business practice 14	One if firm records all sales and purchases, zero otherwise
Business practice 15	One if the firm, using his records, knows how much cash it has on hand?, zero otherwise
Business practice 16	One if it uses product sales information to know if it is growing or decreasing, zero otherwise
Business practice 17	One if it knows the cost of each product it sells, zero otherwise
Business practice 18	One if it knows what products or services the greater utility or benefit, zero otherwise
Business practice 19	One if it keeps a written budget that informs it how much does it owe monthly, zero
Business practice 20	One if the firm has accounting records that document that the business generates sufficient profits to pay a hypothetical bank loan, zero otherwise
Business practice 21	One if the firm saves payment receipts and/or invoices for his suppliers, zero otherwise
Business practice 22	One if it gives payment receipts and or invoices to his customers, zero otherwise
Financial planning	
Business practice 23	One if firm reviews his achievements or financial performance of his business, zero
Business practice 24	One if firm analyzes the areas or activities of the business that can be improved in their performance, zero
Business practice 25	One if firm keeps balance sheet of your business, zero
Business practice 26	One if firm keeps cash flow statements (record of cash available), zero
Communications	
Business practice 27	One if firm discuss business ideas with other people, zero otherwise
Business practice 28	One if it discuss new production techniques with other entrepreneurs, zero otherwise
business practice 29	One if firm meets with at least one networking communities, zero otherwise
Business practice 30	One if firm belongs to an association of entrepreneurs, zero otherwise

Table A3 Cognitive Variables

Variable	Description
Reflective, Intuitive	
Reflective-Intuitive 1	A hamburger and a soda cost 11,000 (COP). The hamburger costs 10,000 (COP) more than the soda. How much does the soda cost? If the answer is 500 it is reflective if it is 1000 it is intuitive
Reflective-Intuitive 2	If it takes 5 machines 5 minutes to make 5 screws, how long does it take 100 machines to make 100 screws? If the answer is 100 min it is reflective if it is 10 min intuitive
Reflective-Intuitive 3	In a lake there is an area covered with floating flowers. Every day the area covered by the flower's doubles in size. If it takes 64 days for the flowers to cover the entire lake, how long does it take them to cover half the lake? If the answer is 63 days it is reflective if it is 32 it is intuitive
Financial mathematical skills	
Mathematical skill 1	In a sales rebate, a store or business sells all of its products at half its price. Before the sale, one of their products cost 4,000,000. At how much will you sell it in the sales rebate?
Mathematical skill 2	If you sold two products for 8,000 each and your customer gave you a 20,000 bill, how much money should you give back?
Mathematical skill 3	Now, assume you have 1,000,000 invested in a business that gives you a 2% profit rate. After 5 years, how much would you have:
Mathematical skill 4	With an annual interest rate of 1% and annual inflation of 2%, how much could you buy the following year?
Perseverance	
Perseverance 1	Many times, I persisted with work when others gave up
Perseverance 2	I keep working on difficult projects even when others object
Personal initiative	
Personal initiative 1	I actively face the problems that come my way
Personal initiative 2	When something goes wrong, I look for a solution immediately
Personal initiative 3	When an opportunity to get involved in something appears, I take it
Personal initiative 4	I take the initiative immediately even when others don't
Personal initiative 5	I take advantages of opportunities quickly to achieve my goals
Personal initiative 6	I usually do more than I am asked
Personal initiative 7	I am particularly good at noticing opportunities

B Summary statistics

Table B1 Descriptive statistics

	mean	sd
Panel A. Socio-demographic		
Female	0.51	0.50
Age 18-30	0.20	0.40
Age 31-42	0.31	0.46
Age 43-58	0.35	0.48
Age more than 58	0.13	0.33
Primary	0.24	0.43
High School	0.35	0.48
College	0.41	0.49
Number of workers	1.41	2.06
Panel B. Economic activity		
Store	0.19	0.39
Prepared food, bars	0.20	0.40
Service (hairdressing,health, etc)	0.23	0.42
Other businesses	0.38	0.49
Panel C. Cities		
Bello	0.09	0.29
Barranquilla	0.12	0.33
Bogota	0.12	0.33
Girardot	0.09	0.29
Soacha	0.13	0.33
Zipaquira	0.08	0.27
Neiva	0.11	0.31
Pereira	0.05	0.22
Bucaramanga	0.13	0.34
Ibague	0.06	0.25
Panel D. Cognitive indicators		
CRT: Reflective	0.21	0.19
CRT: Intuitive	0.61	0.21
Financial mathematical skills	0.58	0.24
Perseverance	0.82	0.20
Panel E. Formality beliefs		
Difference beliefs	0.27	0.35
Observations	1542	

Table B2 Comparative Summary Statistics

	(ENET Survey 2019)		(DANE Survey 2016)		(ANIF Survey 2020)	
	mean	sd	mean	sd	mean	sd
Tax registry	0.84	0.35	0.78	0.41	0.88	n.a.
Commercial Registry	0.76	0.43	0.73	0.45	0.74	n.a.
Bank Loan	0.17	0.38	0.21	0.41	0.27	n.a.
Use electronic wallet	0.13	0.34	0.08	0.28		n.a.
Number of workers	1.42	2.05	2.20	1.53		n.a.
Number of Workers 1					0.32	n.a.
Workers between 2 and 5					0.57	n.a.
Workers between 2 and 5					0.08	n.a.
Observations	1572		33013		1500	

Source: ENET study, DANE and ANIF, data of ANIF survey was extracted from the official report

C Model estimates

C.1 Detailed result tables

Table C1: SEM regressions coefficient estimates (model with controls)

Dependent / Regressor	Estimate	Std.Err	P-val
Business Practices (BP)			
Personal Initiative (PI)	0.06	0.01	<0.001
Female	-0.012	0.007	0.076
High School	0.026	0.01	0.007
College	0.063	0.011	<0.001
Age 31-42	0.005	0.01	0.612
Age 43-58	-0.018	0.01	0.056
Age more than 58	-0.049	0.014	<0.001
Number of workers	0.011	0.002	<0.001
Owner before	0.002	0.011	0.835
Risk aversion	0.008	0.003	0.007
CRT: Reflective	0.083	0.022	<0.001
CRT: Intuitive	0.064	0.018	<0.001
Financial mathematical skills	0.061	0.017	<0.001
Perseverance	0.007	0.019	0.698
Economic activity			
Store (ommitted)			
Prepared food, bars	0.041	0.012	0.001
Service, hairdressing	0.032	0.012	0.005
Other business	0.046	0.011	<0.001
Cities			
Bello (ommitted)			
Barranquilla	0.088	0.018	<0.001
Bogota	-0.006	0.016	0.735
Girardot	-0.109	0.019	<0.001
Soacha	0.003	0.015	0.837
Zipaquira	-0.005	0.018	0.788
Neiva	-0.011	0.017	0.489
Pereira	-0.085	0.02	<0.001
Bucaramanga	-0.045	0.016	0.006
Ibague	-0.019	0.018	0.274
Formality (FO)			
Difference on formality beliefs (DFP)	0.155	0.017	<0.001
Business Practices (BP)	0.071	0.179	0.69
Female	-0.008	0.009	0.337
High School	0.042	0.013	0.001

Continued on next page

Table C1: SEM regressions coefficient estimates (model with controls)

Dependent / Regressor	Estimate	Std.Err	P-val
College	0.063	0.018	0.001
Age 31-42	0.019	0.012	0.12
Age 43-58	0.02	0.013	0.109
Age more than 58	0.006	0.018	0.745
Number of workers	0.009	0.003	0.002
Owner before	0.017	0.014	0.206
Risk aversion	0.008	0.004	0.03
CRT: Reflective	-0.044	0.03	0.14
CRT: Intuitive	-0.016	0.025	0.533
Financial mathematical skills	0.017	0.024	0.489
Perseverance	-0.024	0.024	0.323
Economic activity			
Store (ommitted)			
Prepared food, bars	-0.015	0.017	0.363
Service, hairdressing	0.062	0.017	<0.001
Other business	0.011	0.015	0.471
Cities			
Bello (ommitted)			
Barranquilla	0.053	0.027	0.046
Bogota	0.058	0.021	0.007
Girardot	0.126	0.03	<0.001
Soacha	0.087	0.021	<0.001
Zipaquira	0.115	0.025	<0.001
Neiva	0.115	0.023	<0.001
Pereira	0.02	0.028	0.476
Bucaramanga	0.057	0.023	0.013
Ibague	-0.136	0.025	<0.001
Financial Inclusion (FI)			
Business practices (BP)	1.053	0.346	0.002
Formality (FO)	0.229	0.166	0.167
Female	-0.013	0.016	0.422
High School	0.033	0.025	0.191
College	0.055	0.034	0.11
Age 31-42	0.008	0.022	0.729
Age 43-58	0.005	0.023	0.824
Age more than 58	0.062	0.034	0.071
Number of workers	0.011	0.006	0.053
Owner before	-0.01	0.026	0.705
Risk aversion	-0.002	0.007	0.734
CRT: Reflective	-0.097	0.055	0.08
CRT: Intuitive	-0.017	0.047	0.71
Financial mathematical skills	-0.038	0.045	0.401
Perseverance	0.008	0.046	0.87
Economic activity			
Store (ommitted)			
Prepared food, bars	-0.07	0.031	0.025
Service, hairdressing 3	-0.026	0.029	0.376
Other business	-0.044	0.029	0.127
Cities			
Bello (ommitted)			
Barranquilla	0.189	0.05	<0.001
Bogota	0.05	0.041	0.219
Girardot	0.092	0.059	0.121
Soacha	0.045	0.041	0.267
Zipaquira	-0.026	0.048	0.585
Neiva	0.007	0.046	0.881
Pereira	0.145	0.052	0.006
Bucaramanga	0.072	0.044	0.104
Ibague	0.128	0.049	0.01
Variances			
Financial Inclusion	0.118	0.005	<0.001
Formality	0.088	0.004	<0.001
Business Practices	0.157	0.006	<0.001
Personal Initiative	0.065	0.004	<0.001

Continued on next page

Table C1: SEM regressions coefficient estimates (model with controls)

Dependent / Regressor	Estimate	Std.Err	P-val
Covariances			
Formality ∼ Business Practices	0.004	0.003	0.157
Financial Inclusion ∼ Formality	-0.003	0.004	0.413
Financial Inclusion ∼ Business Practices	-0.01	0.005	0.062

Notes: Coefficients of the regressions of the structural model which was jointly estimated via maximum likelihood. It was implemented using lavaan for R. The description of each measure is available in appendix. A.

C.2 Ordinary least squares and Two-stage least squares

Our empirical strategy is based on the estimation of a SEM. Yet, an alternative is to estimate the parameters of BP and FO in the FI Equation (1) via two-stage least squares (2SLS). The advantage of this alternative is that it requires weaker assumptions on the distribution of the unobservables. It also allows for a simple comparison with the ordinary least squares (OLS) version that imposes exogeneity of BP, FI, and FO.

In this strategy, the equation for BP 2 and a simplified version of the equation for FO 3 became the first stage and the equation for FI 1 the second stage regression. Equation 3 is simplified by removing BP as a regressor, given that our results section has already ruled out this channel. For these regressions, we use the predicted factors after estimating the CFA measurement model.

Table C2 shows the coefficients of the ordinary least squares (OLS) regression (column 1), the first stage regressions for BP (column 2) and FO (column 3), and the 2SLS regression (column 3). All regressions include the same controls as the main exercises, fixed location (city or municipality), and the economic sector fixed effects. Robust standard errors are shown in parentheses.

In the BP first stage (column 2), the PI index has a positive and statistically significant effect on BP. Similarly for the FO equation, DFB is also significant. If we explore the Cragg-Donald F-statistics for both equations (notoriously large based on the rule of thumb in Staiger and Stock (1997)), as well as the weak identification test, we can assert the relevance of both instruments.

Once we address the endogeneity issue (present in the OLS regression in column 1), we find that the estimated effect of BP is much larger than in the OLS method. Therefore, the OLS model underestimates the effect of BP on FI due to the reverse causality issue described above. Furthermore, the coefficient of formality is smaller.

Table C2 Econometric results under OLS and 2SLS

	(1) (OLS) Financial inclusion	(2) (IV 1st stage) Business practices	(3) (IV 1st stage) Formal	(4) (IV 2nd stage) Financial inclusion
Business practices	0.423*** (0.0248)			1.014*** (0.121)
Formal	0.224*** (0.0215)			0.0684 (0.0766)
Personal initiative	0.0421*** (0.00713)	0.0728*** (0.00809)	0.00629 (0.00749)	
Difference formal beliefs	-0.000463 (0.0109)	0.0440*** (0.0113)	0.170*** (0.0116)	
Female	-0.0209*** (0.00588)	-0.0109* (0.00599)	-0.00891 (0.00588)	-0.0158** (0.00681)
High School	0.0491*** (0.00662)	0.0236*** (0.00875)	0.0432*** (0.00863)	0.0418*** (0.00921)
College	0.0948*** (0.00795)	0.0596*** (0.00921)	0.0660*** (0.00923)	0.0698*** (0.0121)
Age 31-42	0.0110 (0.00880)	0.00457 (0.00799)	0.0190** (0.00852)	0.0112 (0.00986)
Age 43-58	-0.00641 (0.00880)	-0.0196** (0.00801)	0.0184** (0.00856)	0.00805 (0.0101)
Age more than 58	0.0310*** (0.0110)	-0.0498*** (0.0116)	0.00229 (0.0117)	0.0608*** (0.0139)
Number of workers	0.0183*** (0.00186)	0.0103*** (0.00175)	0.01000*** (0.00145)	0.0138*** (0.00228)
Owner before	-0.00857 (0.00925)	0.00104 (0.00946)	0.0173* (0.00953)	-0.00650 (0.0101)
Risk aversion	0.00257 (0.00238)	0.00748*** (0.00251)	0.00837*** (0.00251)	-0.000554 (0.00279)
CRT: Reflective	-0.0443** (0.0176)	0.0776*** (0.0171)	-0.0400** (0.0189)	-0.0964*** (0.0222)
CRT: Intuitive	0.0225 (0.0153)	0.0578*** (0.0153)	-0.0130 (0.0162)	-0.0137 (0.0186)
Financial mathematical skills	0.0000995 (0.0138)	0.0524*** (0.0148)	0.0185 (0.0148)	-0.0280 (0.0178)
Perseverance	0.00843 (0.0151)	0.00235 (0.0169)	-0.0234 (0.0165)	0.00339 (0.0189)
Constant	0.00330 (0.0217)	0.00948 (0.0227)	0.00142 (0.0238)	-0.00208 (0.0253)
Observations	1542	1542	1542	1542
Adjusted R^2	0.666	0.405	0.497	0.564
F	83.33	50.04	61.28	58.40
F test of excluded instruments		52.23	108.45	
Underidentification test		91.24	190.87	
Weak identification test		89.58	187.40	

Notes: Standard errors in parentheses, OLS and IV regressions which includes fixed effects of municipality. *
$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D Measurement model

In our study, we aim to explain the impact of BP on FI. and explore the potential channel of FO. However, these concepts are subject to multiple definitions and indicators. For this reason, we construct factors that summarize multiple measures. There are several alternatives to attain this objective. We explore three methods: (i) simple averages, (ii) principal component analysis, and (iii) confirmatory factor analysis.

The first strategy is to add them up and normalize the sum. This provides equal weights to all the measures and has no underlying assumption on what the resulting index means.

The second strategy, PCA, aims to reduce the dimensionality of a dataset by extracting the common variance between the measures. The principal components are variables that group a set of measures according to their correlation and aim to maintain the total variation of the dataset. As a result, the PCA has no 'theory' behind the meaning of the resulting factors and the researcher chooses an interpretation. The reduction of the dimensionality depends on the degree of correlation of the measures. As discussed in the main text, Tables D1 to D3 present the correlation between measures. These correlations are positive and significant in most cases but are not high. As a result, the PCA exercise suggests retaining more than one factor per area of interest.

The third strategy, CFA, needs stronger assumptions in terms of the meaning of the constructs. A set of dedicated measures reflects the behavior of a latent variable plus measurement error. Typically this set of measures is chosen after a PCA exercise (exploratory factor analysis) which guarantees the high correlation (or internal validity) of the measures. As the PCA suggested the usage of more than one factor per dimension, our resulting factors have low reliability. For our exercise, this is not crucial as we are not aiming to assess the quality of a specific instrument that measures an abstract concept (apart from the case of PI, which we build based on the psychology literature). Nevertheless, in section E we try an alternative formulation of the set of measures, to show that our results are valid under alternative definitions of the concepts.

For our main estimates, we prefer the CFA as it is estimated via maximum likelihood as the SEM equations. Therefore, we can avoid the predicting error resulting from estimating a measurement system and then deriving the factors. Moreover, the resulting factor loadings are not very different among the three alternatives, so just presenting one of them is enough.

Table D1 Correlation matrix financial inclusion (FI) measures

	Separate Account	Bank Loan	Family Insurance	Use electronic wallet
Separate Account	1.00			
Bank Loan	-0.01	1.00		
Family Insurance	0.26***	0.07***	1.00	
Use electronic wallet	0.30***	-0.01	0.19***	1.00

Note: Correlation of each components of formality index.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D2 Correlation matrix formality (FO) measures

	Accounting Records	Commercial Registry	Insured Workers	Tax registry	Operating Permit
Accounting records	1.00				
Commercial registry	0.30***	1.00			
Insured workers	0.20***	0.11***	1.00		
Tax registry	0.24***	0.63***	0.13***	1.00	
Operating permit	0.08***	0.34***	0.05	0.27***	1.00

Note: Correlation of each components of formality index.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D3 Correlation matrix business practices (BP) measures

	BP1	BP2	BP3	BP4	BP5	BP6	BP7	BP8
BP1	1							
BP2	0.7901 ***	1						
BP3	0.3335 ***	0.3614 ***	1					
BP4	0.2902 ***	0.3321 ***	0.4175 ***	1				
BP5	0.2144 ***	0.2316 ***	0.2873 ***	0.2472 ***	1			
BP6	0.2089 ***	0.2154 ***	0.2276 ***	0.2017 ***	0.4603 ***	1		
BP7	0.1942 ***	0.1979 ***	0.3360 ***	0.3204 ***	0.3286 ***	0.3441 ***	1	
BP8	0.1852 ***	0.2337 ***	0.2150 ***	0.2174 ***	0.2581 ***	0.2993 ***	0.2417 ***	1
BP9	0.1926 ***	0.1914 ***	0.1989 ***	0.1734 ***	0.2640 ***	0.2861 ***	0.2108 ***	0.2319 ***
BP10	0.1864 ***	0.1847 ***	0.2224 ***	0.3252 ***	0.2149 ***	0.1331 ***	0.2439 ***	0.1262 ***
BP11	0.1903 ***	0.2252 ***	0.2672 ***	0.3544 ***	0.1920 ***	0.1632 ***	0.2186 ***	0.2109 ***
BP12	0.1603 ***	0.1697 ***	0.2043 ***	0.1823 ***	0.2743 ***	0.2818 ***	0.2264 ***	0.3773 ***
BP13	0.1270 ***	0.1606 ***	0.1589 ***	0.1787 ***	0.2532 ***	0.2878 ***	0.1653 ***	0.3885 ***
BP14	0.1330 ***	0.1455 ***	0.1762 ***	0.1352 ***	0.2397 ***	0.2808 ***	0.1731 ***	0.3446 ***
BP15	0.1555 ***	0.1744 ***	0.1896 ***	0.1703 ***	0.1944 ***	0.2555 ***	0.1470 ***	0.3717 ***
BP16	0.1311 ***	0.1699 ***	0.2174 ***	0.2096 ***	0.2497 ***	0.2420 ***	0.1965 ***	0.3911 ***
BP17	-0.0015	0.0353	0.0867 ***	0.1092 ***	0.048	0.0971 ***	0.1347 ***	0.1175 ***
BP18	0.0409	0.0507 ***	0.1458 ***	0.0933 ***	0.0593 ***	0.1311 ***	0.1806 ***	0.2075 ***
BP19	0.1504 ***	0.2030 ***	0.2130 ***	0.1717 ***	0.2676 ***	0.2004 ***		0.3370 ***
BP20	0.1569 ***	0.1718 ***	0.1931 ***	0.1474 ***	0.2345 ***	0.2354 ***	0.1763 ***	0.2085 ***
BP21	0.0692 ***	0.1004 ***	0.1221 ***	0.1740 ***	0.1475 ***	0.1561 ***	0.1195 ***	0.2294 ***
BP22	0.1467 ***	0.1534 ***	0.1168 ***	0.0800 ***	0.1939 ***	0.2308 ***	0.1557 ***	0.1835 ***
BP23	0.1657 ***	0.1863 ***	0.1891 ***	0.1359 ***	0.2377 ***	0.2623 ***	0.1947 ***	0.4302 ***
BP24	0.1340 ***	0.1887 ***	0.1869 ***	0.1839 ***	0.1989 ***	0.2521 ***	0.2444 ***	0.3946 ***
BP25	0.1286 ***	0.1406 ***	0.1674 ***	0.1025 ***	0.2623 ***	0.2670 ***	0.1995 ***	0.3412 ***
BP26	0.1502 ***	0.1650 ***	0.1876 ***	0.1164 ***	0.2630 ***	0.2852 ***	0.1925 ***	0.3009 ***
BP27	0.2755 ***	0.2770	0.2868 ***	0.2409 ***	0.2448 ***	0.2598 ***	0.2186 ***	0.2436 ***
BP28	0.2393 ***	0.2617 ***	0.2584 ***	0.2166 ***	0.2990 ***	0.2990 ***	0.2253 ***	0.2256 ***
BP29	0.1171 ***	0.1072 ***	0.0958 ***	0.1025 ***	0.1190 ***	0.1376 ***	0.0751 ***	0.0560
BP30	0.1504 ***	0.1311 ***	0.0675 ***	0.0634 ***	0.1039 ***	0.0663 ***	0.0974 ***	0.0493

	BP9	BP10	BP11	BP12	BP13	BP14	BP15	BP16
BP9	1							
BP10	0.2399 ***	1						
BP11	0.2051 ***	0.4893 ***	1					
BP12	0.2545 ***	0.2384 ***	0.2572 ***	1				
BP13	0.2235 ***	0.1694 ***	0.2448 ***	0.6797 ***	1			
BP14	0.2416 ***	0.1134 ***	0.1575 ***	0.4937 ***	0.5175 ***	1		
BP15	0.1988 ***	0.1215 ***	0.1901 ***	0.4884 ***	0.5007 ***	0.5493 ***	1	
BP16	0.2484 ***	0.1466 ***	0.2531 ***	0.4714 ***	0.4582 ***	0.4943 ***	0.5384 ***	1
BP17	0.0197	0.0728 ***	0.1840 ***	0.1316 ***	0.1509 ***	0.1262 ***	0.1988 ***	0.2089 ***
BP18	0.0621 ***	0.0756 ***	0.1544 ***	0.2117 ***	0.2037 ***	0.1764 ***	0.2649 ***	0.3165 ***
BP19	0.2763 ***	0.1305 ***	0.1868 ***	0.4084 ***	0.4328 ***	0.4318 ***	0.4295 ***	0.4505 ***
BP20	0.2724 ***	0.1988 ***	0.1871 ***	0.3759 ***	0.4068 ***	0.3606 ***	0.3143 ***	0.3130 ***
BP21	0.1721 ***	0.1617 ***	0.2289 ***	0.2975 ***	0.2846 ***	0.3099 ***	0.3196 ***	0.2968 ***
BP22	0.2810 ***	0.1029 ***	0.1295 ***	0.2398 ***	0.2556 ***	0.3167 ***	0.2534 ***	0.3008 ***
BP23	0.2627 ***	0.1391 ***	0.2107 ***	0.4513 ***	0.4565 ***	0.4399 ***	0.4709 ***	0.5101 ***
BP24	0.2254 ***	0.1756 ***	0.2757 ***	0.3935 ***	0.3906 ***	0.3473 ***	0.3986 ***	0.4578 ***
BP25	0.2132 ***	0.1102 ***	0.1647 ***	0.4746 ***	0.5257 ***	0.4751 ***	0.4218 ***	0.4561 ***
BP26	0.2689 ***	0.1068 ***	0.1130 ***	0.4016 ***	0.4688 ***	0.4117 ***	0.3911 ***	0.3871 ***
BP27	0.3124 ***	0.2089 ***	0.2751 ***	0.2576 ***	0.2563 ***	0.2225 ***	0.1985 ***	0.2639 ***
BP28	0.3422 ***	0.1626 ***	0.2150 ***	0.2674 ***	0.3001 ***	0.2619 ***	0.2054 ***	0.2658 ***
BP29	0.2166 ***	0.0668 ***	0.0444	0.1062 ***	0.1167 ***	0.1041 ***	0.0787 ***	0.0792 ***
BP30	0.1880 ***	0.1119 ***	0.0376	0.1184 ***	0.1197 ***	0.0913 ***	0.0913 ***	0.0767 ***

	BP17	BP18	BP19	BP20	BP21	BP22	BP23	BP24
BP17	1							
BP18	0.4249 ***	1						
BP19	0.1873 ***	0.2089 ***	1					
BP20	0.0499	0.1349 ***	0.4403 ***	1				
BP21	0.2595 ***	0.1650 ***	0.2359 ***	0.1771 ***	1			
BP22	0.0481	0.0576 ***	0.3099 ***	0.3460 ***	0.1626 ***	1		
BP23	0.1215 ***	0.2060 ***	0.3802 ***	0.3654 ***	0.2013 ***	0.3371 ***	1	
BP24	0.1836 ***	0.2950 ***	0.3500 ***	0.3086 ***	0.2425 ***	0.2526 ***	0.5535 ***	1
BP25	0.1513 ***	0.1655 ***	0.4401 ***	0.4432 ***	0.2359 ***	0.3598 ***	0.4996 ***	0.4087 ***
BP26	0.0967 ***	0.1473 ***	0.4492 ***	0.5340 ***	0.2082 ***	0.3841 ***	0.4189 ***	0.3562 ***
BP27	0.0820 ***	0.1244 ***	0.2445 ***	0.3005 ***	0.1269 ***	0.2338 ***	0.2689 ***	0.3055 ***
BP28	0.0501 ***	0.1041 ***	0.2964 ***	0.3484 ***	0.1218 ***	0.2777 ***	0.2847 ***	0.2667 ***
BP29	-0.0227	-0.0185	0.1397 ***	0.1682 ***	0.0322	0.1799 ***	0.1344 ***	0.0853 ***
BP30	-0.0302	0.0073	0.1379 ***	0.1584 ***	0.0272	0.1515 ***	0.1369 ***	0.1030 ***

	BP25	BP26	BP27	BP28	BP29	BP30
BP25	1					
BP26	0.5919 ***	1				
BP27	0.2615 ***	0.2987 ***	1			
BP28	0.2903 ***	0.3629 ***	0.6136 ***	1		
BP29	0.1303 ***	0.1731 ***	0.2560 ***	0.3267 ***	1	
BP30	0.1706 ***	0.1769 ***	0.1947 ***	0.2462 ***	0.3830 ***	1

Note: Correlation of each components of formality index.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D.1 Simple averages

Under this alternative, the indices are simple averages of their inputs:

$$Index_i = \frac{1}{J} \sum_{j=1}^{J} variable_{ij} \tag{4}$$

Where, $variable_{ij}$ corresponds to the value of variable j of the J variables that are part of the particular index, for a given business i. The value of each index ranges from 0 to 1.

Table D4 presents the main SEM model estimates using the resulting indices. While coefficients are different (as the domain of the indices differs), qualitative results are the same as those presented in Table 2.

Table D4: Main Results using simple averages

	(1) No controls	(2) Controls
Panel A. Main estimated coefficients		
β_3: Personal Initiative → Business practices	0.322***	0.294***
	(0.042)	(0.039)
β_5: Business practices → Formal	0.195	0.171
	(0.136)	(0.148)
β_4: Difference formal beliefs → Formal	0.213***	0.173***
	(0.017)	(0.018)
β_2: Formal → Financial Inclusion	0.079	0.049
	(0.071)	(0.101)
Panel B. Paths from BP to FI		
$\beta_1 + \beta_2 \cdot \beta_5$: Business practices $\overrightarrow{Total}$ Financial Inclusion	0.516***	0.459***
	(0.125)	(0.146)
	[100%]	[100%]
β_1: Business practices $\overrightarrow{Direct}$ Financial Inclusion	0.501***	0.451***
	(0.125)	(0.148)
	[97.1%]	[98.2%]
$\beta_2 \cdot \beta_5$: Business practices → Formal → Financial Inclusion	0.015	0.008
	(0.017)	(0.019)
	[2.9%]	[1.8%]
Observations	1542	1542
RMSEA	0.143	0.111
SRMR	0.036	0.004
P-value (Chi-square)	0.000	0.000
Comparative Fit Index (CFI)	0.952	0.988
Tucker-Lewis Index (TLI)	0.571	-0.006

Notes: The model, is a estimation of the structural equation model via maximum likelihood using Lavaan package for R. The factors were constructed as simple averages of all the relevant measures. Standard errors in parentheses. Percentage of the total effect in brackets, in Panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.2 Principal component analysis

Principal components analysis (PCA), is a multivariate technique introduced by Pearson (1901) used to describe the relationship between several response

variables and to explain the total variation in the data. PCA uses a few equations constructed from the original variables, which are called components. PCA is very useful when the variables under study are highly correlated (positively or negatively) or when the number of independent variables is large.

The first step in this methodology consists in analyzing the correlation matrix for all variables, which shows those correlations are not particularly high (Tables D1 to D3). This indicates that several principal components will be required to account for the variation in the data.

The second step is to execute the PCA and decide the number of factors to retain per concept. Tables D5 to D8 present four eigenvectors (principal components) and the associated eigenvalues and cumulative variance. For FI, these are all the eigenvectors, for FO these four represent 92% of the total variance. However, for BP only 45,8%; we would need 22 factors out of 30 to get 90%, and for PI is 79.4% (6 factors out of 7 to get 90%). If we consider eigenvectors greater than 1 instead (Kaiser's rule), for FI, FO, and PI two would be retained, and more than four for BP.

Rather than generating several factors based on the PCA, in Appendix E we consider each individual measure of FI and FO directly (not aggregating them in an index), and BP in its specific sections.

Finally, as a comparison with the other exercises where only one factor is constructed per concept, Table D9 presents the main results using the first component for each measure. Once again, the magnitude of coefficients is not directly comparable with those in Table 2, but the qualitative results are the same.

Table D5 Financial Inclusion PCA

	Comp1	Comp2	Comp3	Comp4
Separate Account	.6180278	-.0952988	.0113874	-.7802757
Bank Loan	.0573795	.9480146	.3060013	-.0658715
Family Insurance	.548789	.2208285	-.7014878	.3974673
Use electronic wallet	.559982	-.2083774	.6435436	.4783833
Eigenvalue	1.5011	1.0213	.7922	.6853
Cumulative	0.3753	0.6306	0.8287	1.0000

Table D6 Formality PCA

	Comp1	Comp2	Comp3	Comp4
Accounting records	.3729988	.47705	-.4831589	.6262796
Commercial Registry	.5908192	-.1663817	-.1508395	-.2387093
Insured Workers	.2190069	.7399104	.5905858	-.2332912
Tax registry	.5678718	-.1484322	-.1564932	-.4382625
Operating Permit	.3759816	-.4186178	.6087062	.5516276
Eigenvalue	2.0526	1.0500	.8236	.7150
Cumulative	0.4105	0.6205	0.7853	0.9283

Table D7 Business Practices PCA

	Comp1	Comp2	Comp3	Comp4
BP1	.1323776	.3628865	.0102176	-.4869251
BP2	.1453354	.3572742	.044645	-.4799822
BP3	.1520983	.2784342	.1500778	-.0709173
BP4	.1392047	.297759	.222587	.0456015
BP5	.1651439	.167413	-.0027428	.0275488
BP6	.1757686	.106371	-.009732	.0244929
BP7	.1495485	.1942835	.1393311	.1551535
BP8	.1999424	-.0414327	.1162962	-.1282857
BP9	.1664052	.1289929	-.1500544	.1681628
BP10	.1241117	.2251013	.1747657	.2919771
BP11	.1517745	.1809174	.2636844	.2594972
BP12	.2387304	-.1438434	.0394501	-.0466023
BP13	.2414022	-.1843938	-.0028302	-.0600831
BP14	.2264371	-.1922733	-.0196934	-.1165329
BP15	.2238214	-.2012311	.0853191	-.1445619
BP16	.2357759	-.1671212	.102682	-.0397397
BP17	.0887436	-.0990799	.3458688	.222741
BP18	.1175112	-.1136569	.3090727	.1778896
BP19	.222573	-.1170903	-.0452966	-.04466
BP20	.2078919	-.0599351	-.1923322	.0334844
BP21	.1460039	-.0988168	.1976035	.0774725
BP22	.1681325	-.0464082	-.2235231	.0082528
BP23	.2344761	-.1469014	-.0142535	-.0818838
BP24	.2171722	-.0926273	.0970005	.036607
BP25	.234351	-.1917674	-.1147752	-.0668448
BP26	.2280044	-.1315284	-.2069802	-.0433199
BP27	.1852962	.2014193	-.1536709	.1955486
BP28	.1923925	.1692348	-.2514025	.1972178
BP29	.0898277	.1327797	-.3747694	.2295528
BP30	.086489	.1119015	-.3381261	.1617214
Eigenvalue	8.1682	2.4221	1.8416	1.3244
Cumulative	0.2723	0.3530	0.4144	0.4586

Table D8 Personal Initiative PCA

	Comp1	Comp2	Comp3	Comp4
PI1	.3457649	.5778131	.2560615	.2808361
PI2	.3658673	.5644485	.1032182	-.0615997
PI3	.3989678	.0146333	-.43586	-.4826584
PI4	.4219895	-.2140715	-.0770025	-.042738
PI5	.4256914	-.1884027	-.2166217	-.1967557
PI6	.2976476	-.3939014	.805242	-.2273013
PI7	.3735372	-.3329194	-.1803763	.7695273
Eigenvalue	3.1404	1.0625	.7647	.5951
Cumulative	0.4486	0.6004	0.7097	0.7947

Table D9: Main Results using PCA

	(1) No controls	(2) Controls
Panel A. Main estimated coefficients		
β_3: Personal Initiative → Business practices	0.295***	0.268***
	(0.040)	(0.037)
β_5: Business practices → Formal	0.056	0.050
	(0.063)	(0.068)
β_4: Difference formal beliefs → Formal	1.233***	1.029***
	(0.092)	(0.098)
β_2: Formal → Financial Inclusion	0.098	0.089
	(0.067)	(0.092)
Panel B. Paths from BP to FI		
$\beta_1 + \beta_2 \cdot \beta_5$: Business practices $\overrightarrow{Total}$ Financial Inclusion	0.226***	0.218***
	(0.057)	(0.067)
	[100%]	[100%]
β_1: Business practices $\overrightarrow{Direct}$ Financial Inclusion	0.221***	0.213***
	(0.057)	(0.068)
	[97.8%]	[97.7%]
$\beta_2 \cdot \beta_5$: Business practices → Formal → Financial Inclusion	0.005	0.004
	(0.007)	(0.008)
	[2.2%]	[2.3%]
Observations	1542	1542
RMSEA	0.149	0.113
SRMR	0.039	0.004
P-value (Chi-square)	0.000	0.000
Comparative Fit Index (CFI)	0.951	0.989
Tucker-Lewis Index (TLI)	0.560	0.048

Notes: The model, is a estimation of the structural equation model via maximum likelihood using Lavaan package for R. The factors correspond to the first principal component of each set of measures. Standard errors in parentheses. Percentage of the total effect in brackets, in Panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.3 Confirmatory Factor Analysis

The standardized version of each of these measures is denoted as $M_i^{m(j)}$, and is assumed to be part of a set $\mathcal{M}^j$ that noisily captures information from a factor F_i^j where $j \in \{FO, FI, BP, PI\}$ for microbusiness i. It is also assumed that the noise $\phi_i^{m(j)}$, a classical measurement error[17], is additively separated from the factor that is loaded into the measure by the parameter $\varphi^{m(j)}$. This is presented in Equation 5, where $v^{m(j)}$ is a measure-specific intercept. Given that the factors are latent variables, identification requires normalization of the location and scale (Anderson and Rubin, 1956; Heckman et al., 2013). One of the factor loadings $\varphi^{m(j)}$ is set to 1 for each set $\mathcal{M}^j$ (scale), and the mean of all factors to 0 (location).

$$M_i^{m(j)} = v^{m(j)} + \varphi^{m(j)} F_i^j + \phi_i^{m(j)}, \quad m(j) \in \mathcal{M}^j, \quad j \in \{FO, FI, BP, PI\} \quad (5)$$

The estimates of these parameters are presented in Table D10 below. On top of the factor loadings, intercept, and variance of each measure (see equation 5 above, the table also presents two reliability measures: (i) the Cronbach's alpha, and (ii) the average variance extracted (AVE). For the case of FI and FO, both the alpha and the AVE are low. As discussed above, the correlation among measures is not extremely high, so the internal consistency of the resulting factor is not the highest. For the case of BP and PI, the alpha coefficient is above 0.70, which indicates internal validity. Yet, the AVE for both of them is below 0.50, indicating that less than 50% of the variance of the factors is explained with the observed measures. These findings motivate alternative exercises with fewer measures per factor, just as the PCA approach. This is presented in appendix E.

In the results section of the main text, we predict the factors from the measurement system. Here, Figure D1 presents the densities of the predicted factors. Moreover, Figure D2 shows that BP and FI correlation is across the entire domain, and follows a linear pattern over most of it.

---

[17] Each of them is assumed to be *iid*, normally distributed with mean 0 and sample variance $\sigma^{m(j)}$.
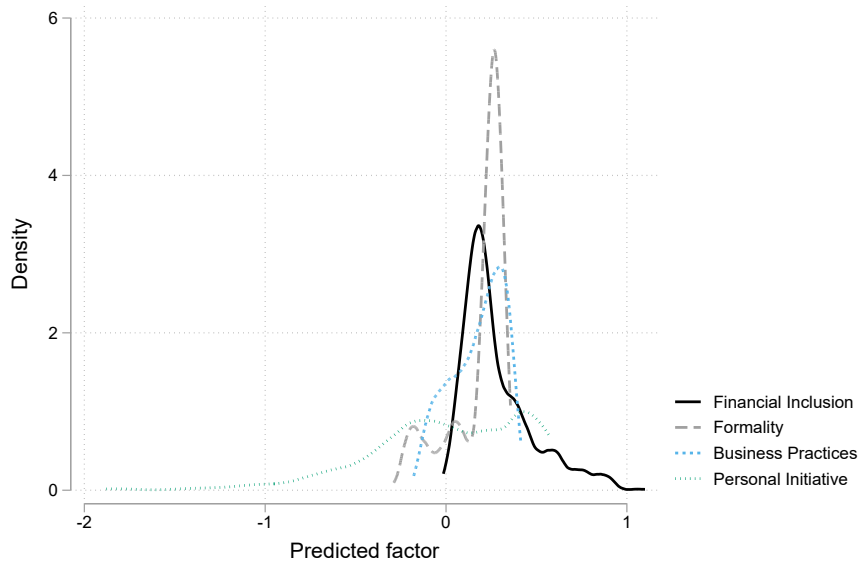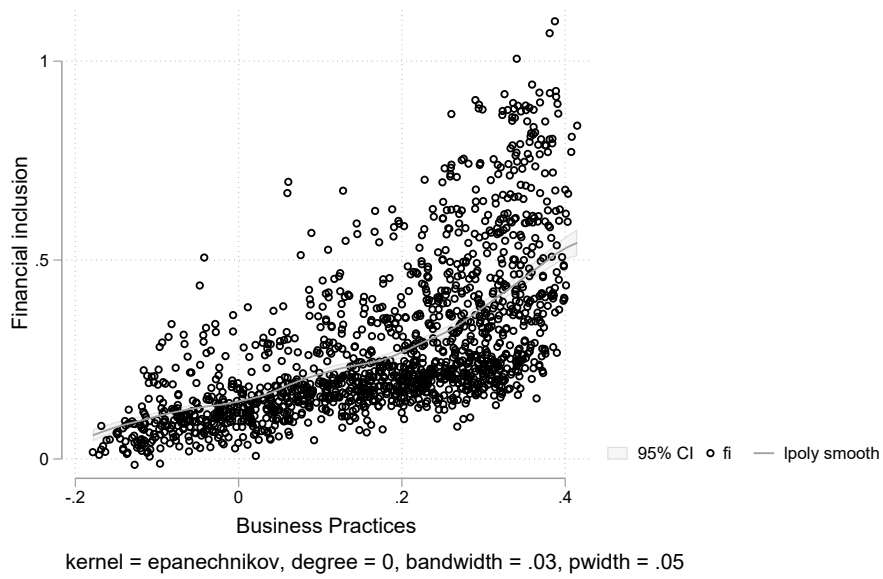
Fig. D1 Distribution of the predicted factors

Fig. D2 Non-linear regression BP and FI

kernel = epanechnikov, degree = 0, bandwidth = .03, pwidth = .05

Table D10 CFA items and estimated coefficients

Factor / Measure	Mean	Factor coefficient ϕ			Other parameters	
		Estimate	Std.Err	P-val	Intercept	Error. Var
Financial Inclusion (FI), Alpha=0.378, AVE=0.203						
Separate Account		1			-0.077	0.089
Use electronic wallet		-0.035	0.05	0.487	0.183	0.144
Bank Loan		0.602	0.056	<0.001	0.013	0.118
Family Insurance		0.685	0.058	<0.001	-0.043	0.088
Formality (FO), Alpha=0.588, AVE=0.285						
Operating Permit		1			0.563	0.157
Accounting records		1.931	0.143	<0.001	0.401	0.065
Commercial Registry		0.515	0.076	<0.001	0.175	0.181
Tax registry		1.513	0.113	<0.001	0.566	0.057
Insured Workers		1.11	0.108	<0.001	0.336	0.21
Business Practices (BP), Alpha=0.904, AVE=0.266						
BP1: Marketing 1		1			0.21	0.213
BP2: Marketing 2		1.116	0.116	<0.001	0.191	0.207
BP3: Marketing 3		1.192	0.124	<0.001	0.318	0.216
BP4: Marketing 4		1.04	0.115	<0.001	0.427	0.213
BP5: Marketing 5		1.39	0.136	<0.001	0.314	0.202
BP6: Marketing 6		1.505	0.142	<0.001	0.326	0.19
BP7: Marketing 7		1.116	0.116	<0.001	0.493	0.186
BP8: Marketing 8		1.709	0.154	<0.001	0.357	0.158
BP9: Marketing 9		1.407	0.136	<0.001	0.159	0.194
BP10: Inventory 1		0.94	0.111	<0.001	0.417	0.223
BP11: Inventory 2		1.136	0.118	<0.001	0.475	0.19
BP12: Inventory 3		2.106	0.181	<0.001	0.294	0.121
BP13: Inventory 4		2.233	0.191	<0.001	0.204	0.125
BP14: Sales and Purchases 1		1.949	0.169	<0.001	0.368	0.119
BP15: Sales and Purchases 2		1.844	0.16	<0.001	0.42	0.113
BP16: Sales and Purchases 3		1.995	0.172	<0.001	0.359	0.115
BP17: Sales and Purchases 4		0.376	0.053	<0.001	0.864	0.062
BP18: Sales and Purchases 5		0.576	0.066	<0.001	0.813	0.072
BP19: Sales and Purchases 6		1.997	0.174	<0.001	0.246	0.148
BP20: Sales and Purchases 7		1.868	0.166	<0.001	0.116	0.165
BP21: Sales and Purchases 8		0.896	0.091	<0.001	0.705	0.102
BP22: Sales and Purchases 9		1.528	0.145	<0.001	0.199	0.194
BP23: Financial planning 1		2.091	0.18	<0.001	0.289	0.125
BP24: Financial planning 2		1.761	0.155	<0.001	0.423	0.126
BP25: Financial planning 3		2.143	0.184	<0.001	0.245	0.129
BP26: Financial planning 4		2.111	0.183	<0.001	0.095	0.144
BP27: Communications 1		1.549	0.144	<0.001	0.145	0.186
BP28: Communications 2		1.512	0.139	<0.001	0.033	0.154
BP29: Communications 3		0.417	0.058	<0.001	0.015	0.075
BP30: Communications 4		0.401	0.058	<0.001	0.019	0.077
Personal Initiative (PI), Alpha=0.786, AVE=0.358						
Personal initiative 1		1			4.332	0.697
Personal initiative 2		0.949	0.061	<0.001	4.482	0.521
Personal initiative 3		1.463	0.09	<0.001	3.988	0.696
Personal initiative 4		1.549	0.093	<0.001	4.06	0.562
Personal initiative 5		1.497	0.09	<0.001	4.166	0.487
Personal initiative 6		1.123	0.087	<0.001	3.846	1.161
Personal initiative 7		1.336	0.087	<0.001	3.977	0.731

Notes: Coefficients of the measurement system (a confirmatory factor analysis) of the structural model which was jointly estimated via maximum likelihood. It was implemented using lavaan for R. *Alpha* corresponds to the 'Cronbach's alpha', ad *AVE* to the 'average variance extracted', which are reliability measures obtained by estimating alone the CFA component. The description of each measure is available in appendix A.

Table E11 Econometric results SEM Model with different measures of FI

	(1) Bank Account	(2) Bank loan	(3) Family insurance	(4) Electr wallet
Panel A. Main estimated coefficients				
β_3: Personal Initiative $\rightarrow$ Business practices	0.060*** (0.010)	0.060*** (0.010)	0.060*** (0.010)	0.060*** (0.010)
β_5: Business practices $\rightarrow$ Formal	0.072 (0.179)	0.071 (0.179)	0.071 0.179	0.071 0.178
β_4: Difference formal beliefs $\rightarrow$ Formal	0.155 *** (0.017)	0.155 *** (0.017)	0.155 *** (0.017)	0.155 *** (0.017)
β_2: Formal $\rightarrow$ Financial Inclusion	0.317 (0.199)	-0.122 (0.213)	-0.191 (0.206)	0.243 (0.181)
Panel B. Paths from BP to FI				
$\beta_1 + \beta_2 \cdot \beta_5$: Business practices $\overrightarrow{Total}$ Financial Inclusion	1.256*** (0.407) [100%]	0.477 (0.416) [100%]	0.737* (0.401) [100%]	0.419 (0.348) [100%]
β_1: Business practices $\overrightarrow{Direct}$ Financial Inclusion	1.233*** (0.411) [98.1%]	0.486 (0.418) [101.8%]	0.750* (0.406) [101.7%]	0.402 (0.352) [96%]
$\beta_2 \cdot \beta_5$: Business practices $\rightarrow$ Formal $\rightarrow$ Financial Inclusion	0.023 (0.058) [1.9%]	-0.009 (0.026) [-1.8%]	-0.014 (0.037) [-1.7%]	0.017 (0.045) [4%]
Observations	1542	1542	1542	1542
RMSEA	0.059	0.059	0.059	0.059
SRMR	0.062	0.062	0.061	0.061
P-value (Chi-square)	0.000	0.000	0.000	0.000
Comparative Fit Index (CFI)	0.596	0.593	0.595	0.596
Tucker-Lewis Index (TLI)	0.570	0.567	0.568	0.569

Notes: The model is a estimation of the structural equation model jointly with the measurement system via maximum likelihood using Lavaan package for R. Standard errors in parentheses. Percentage of the total effect in brackets, in Panel B. * (p<0.10), ** (p<0.05), *** (p<0.01)

E Alternative definition of the items

In this section, we conduct several exercises to determine that our results are maintained when considering different measures of FI, BP, and FO. This follows the results from the measurement system suggesting that more than one factor would be necessary to summarize the total variances of the system.

We ran several additional exercises looking at how results using first, each of the four components of the FI construct and testing how BP affects each component of FI, and second, using four out of the five groups of business practices and testing how each of the BP group affects FI. Results are presented in Table E11) and Table E12). Table E13) presents the effects of BP through the specific formality items.

Table E12 Econometric results SEM Model with different measures of BP

	(1) Marketing	(2) Inv-Sales	(3) Fin plan - Communi	(4) Sales-purch
Panel A. Main estimated coefficients				
β_3: Personal Initiative $\rightarrow$ Business practices	0.107*** (0.017)	0.097*** (0.015)	0.072*** (0.015)	0.083*** (0.015)
β_5: Business practices $\rightarrow$ Formal	0.027 (0.093)	0.044 (0.110)	0.049 (0.143)	0.048 (0.127)
β_4: Difference formal beliefs $\rightarrow$ Formal	0.153*** (0.017)	0.155*** (0.017)	0.154*** (0.017)	0.156*** (0.017)
β_2: Formal $\rightarrow$ Financial Inclusion	0.308* (0.168)	0.259 (0.165)	0.251 (0.165)	0.256 (0.165)
Panel B. Paths from BP to FI				
$\beta_1 + \beta_2 \cdot \beta_5$: Business practices $\overrightarrow{Total}$ Financial Inclusion	0.589*** (0.193) [100%]	0.649*** (0.208) [100%]	0.890*** (0.297) [100%]	0.763*** (0.248) [100%]
β_1: Business practices $\overrightarrow{Direct}$ Financial Inclusion	0.580*** (0.193) [98.5%]	0.638*** (0.210) [98.3%]	0.878*** (0.299) [98.6%]	0.751*** (0.250) [98.4%]
$\beta_2 \cdot \beta_5$: Business practices $\rightarrow$ Formal $\rightarrow$ Financial Inclusion	0.008 (0.029) [1.5%]	0.011 (0.029) [1.7%]	0.012 (0.037) [1.4%]	0.012 (0.033) [1.6%]
Observations	1542	1542	1542	1542
RMSEA	0.054	0.057	0.056	0.057
SRMR	0.045	0.046	0.046	0.046
P-value (Chi-square)	0.000	0.000	0.000	0.000
Comparative Fit Index (CFI)	0.704	0.688	0.699	0.688
Tucker-Lewis Index (TLI)	0.643	0.624	0.636	0.624

Notes: The model, is a estimation of the structural equation model jointly with the measurement system via maximum likelihood using Lavaan package for R. Standard errors in parentheses. Percentage of the total effect in brackets, in Panel B. * (p<0.10), ** (p<0.05), *** (p<0.01)

F Further heterogeneous effects

Here we explore two alternative sets of heterogeneous effects. We consider the age of the entrepreneur as well as economic activity. Results here are suggestive, as the reduction on sample size

With respect to age, we find that our results hold in almost all cases except in the case of older (age 58+) owners (Table F14, columns 1 to 4). In particular, PI is not statistically significant for BP, and MP does not affect BP. This could be explained by young managers paying more attention to personal discovery, emphasizing self-motivation and self-discipline (Birkinshaw et al., 2019). In contrast, older managers are more reflective thinkers and place greater weight on learning from setbacks and knowing their strengths. However, the smaller sample size for the 58+ group might be also behind this finding.

Finally, on dividing the sample by economic activity (Table F14, columns 5 to 8), most relationships are qualitatively the same. For food, bars, and services the BP to FI paths is still positive but insignificant.

Table E13 Econometric results SEM Model with different measures of FO

	(1) Accounting records	(2) Comm registry	(3) Tax registry	(4) Insured workers
Panel A. Main estimated coefficients				
β_3: Personal Initiative $\rightarrow$ Business practices	0.060*** (0.010)	0.060*** (0.010)	0.060*** (0.010)	0.060*** (0.010)
β_5: Business practices $\rightarrow$ Formal	-0.146 (0.403)	0.508 (0.439)	0.069 (0.335)	0.215 (0.478)
β_4: Difference formal beliefs $\rightarrow$ Formal	0.329*** (0.032)	0.062*** (0.035)	0.231*** (0.027)	0.216*** (0.039)
β_2: Formal $\rightarrow$ Financial Inclusion	0.108 (0.078)	0.591 (0.500)	0.155 (0.112)	0.166 (0.117)
Panel B. Paths from BP to FI				
$\beta_1 + \beta_2 \cdot \beta_5$: Business practices $\overrightarrow{Total}$ Financial Inclusion	1.066*** (0.343) [100%]	1.092*** (0.349) [100%]	1.064*** (0.343) [100%]	1.058*** (0.342) [100%]
β_1: Business practices $\overrightarrow{Direct}$ Financial Inclusion	1.082*** (0.347) [101.5%]	0.791*** (0.484) [72.4%]	1.053*** (0.348) [98.9%]	1.023*** (0.341) [96.7%]
$\beta_2 \cdot \beta_5$: Business practices $\rightarrow$ Formal $\rightarrow$ Financial Inclusion	-0.016 (0.045) [-1.5%]	0.300 (0.369) [27.6%]	0.011 (0.052) [1.1%]	0.036 (0.084) [3.3%]
Observations	1542	1542	1542	1542
RMSEA	0.057	0.057	0.057	0.057
SRMR	0.054	0.055	0.054	0.055
P-value (Chi-square)	0.000	0.000	0.000	0.000
Comparative Fit Index (CFI)	0.611	0.607	0.610	0.610
Tucker-Lewis Index (TLI)	0.585	0.580	0.584	0.583

Notes: The model, is a estimation of the structural equation model jointly with the measurement system via maximum likelihood using Lavaan package for R. Standard errors in parentheses. Percentage of the total effect in brackets, in Panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table F14 Heterogeneous effects

	(1) (18-30)	(2) (31-42)	(3) (43-58)	(4) (more than 58)	(5) (Store)	(6) (Food, bars)	(7) (Service)	(8) (Other business)
Panel A. Main estimated coefficients								
β_3: Personal Initiative $\rightarrow$ Business practices	0.060*** (0.022)	0.050*** (0.019)	0.081*** (0.021)	0.015 (0.014)	0.049*** (0.017)	0.119*** (0.038)	0.022* (0.012)	0.075*** (0.021)
β_5: Business practices $\rightarrow$ Formal	0.046 (0.401)	-0.018 (0.137)	0.074 (0.379)	1.939 (3.915)	-0.084 (0.441)	-0.023 (0.146)	-1.485 (1.331)	0.319 (0.255)
β_4: Difference formal beliefs $\rightarrow$ Formal	0.055 (0.037)	0.061*** (0.017)	0.279*** (0.039)	0.283*** (0.063)	0.200*** (0.047)	0.095*** (0.031)	0.139*** (0.036)	0.147*** (0.026)
β_2: Formal $\rightarrow$ Financial Inclusion	1.896 (1.616)	0.461 (0.915)	-0.015 (0.142)	0.198 (0.140)	-0.305 (0.226)	0.003 (0.605)	0.728 (0.429)	0.433 (0.310)
Panel B. Paths from BP to FI								
$\beta_1 + \beta_2 \cdot \beta_5$: Business practices $\overrightarrow{Total}$ Financial Inclusion	0.987 (0.680) [100%]	1.974** (0.917) [100%]	0.626 (0.447) [100%]	-3.517 (4.420) [100%]	1.481** (0.721) [100%]	0.620 (0.387) [100%]	0.687 (1.541) [100%]	1.092** (0.531) [100%]
β_1: Business practices $\overrightarrow{Direct}$ Financial Inclusion	0.899 (1.009) [104.3%]	1.982** (0.921) [94.1%]	0.628 (0.448) [94.1%]	-3.900 (4.572) [94.1%]	1.455** (0.716) [94.1%]	0.620 (0.387) [94.1%]	1.767 (1.914) [94.1%]	0.954* (0.541) [94.1%]
$\beta_2 \cdot \beta_5$: Business practices $\rightarrow$ Formal $\rightarrow$ Financial Inclusion	0.088 (0.766) [-4.3%]	-0.008 (0.065) [5.8%]	-0.001 (0.012) [5.8%]	0.383 (0.818) [5.8%]	0.026 (0.135) [5.8%]	-0.000 (0.014) [5.8%]	-1.080 (1.150) [5.8%]	0.138 (0.147) [5.8%]
Observations	317	482	547	196	296	302	356	588
RMSEA	0.058	0.061	0.060	0.070	0.066	0.063	0.061	0.061
SRMR	0.070	0.069	0.066	0.081	0.079	0.076	0.072	0.066
P-value (Chi-square)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Comparative Fit Index (CFI)	0.565	0.553	0.603	0.553	0.538	0.516	0.571	0.576
Tucker-Lewis Index (TLI)	0.538	0.526	0.579	0.526	0.510	0.487	0.544	0.550

Notes: The model, is a estimation of the structural equation model jointly with the measurement system via maximum likelihood using Lavaan package for R. Standard errors in parentheses. Percentage of the total effect in brackets, in Panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$