Online Appendix for

On the Origins of Entrepreneurship: Evidence from Sibling Correlations

Theodor Vladasel* Matthew J. Lindquist† Joeri Sol‡ Mirjam van Praag§

March 2, 2020

Abstract

This online appendix includes additional material for our paper entitled "On the Origins of Entrepreneurship: Evidence from Sibling Correlations", accepted for publication in the *Journal of Business Venturing*. Appendix A formally describes the models used in our empirical analysis, including i) sibling correlations, ii) the relationship between sibling correlations and intergenerational correlations, and iii) extensions designed to assess the contributions of neighborhoods, parental characteristics, sibling peer effects, and shared genes to sibling correlations. Appendix B contains figures and tables that provide additional information, analysis, and robustness checks for our main results.

* Corresponding author. Universitat Pompeu Fabra and Barcelona GSE, theodor.vladasel@upf.edu

† Swedish Institute for Social Research (SOFI), Stockholm University, matthew.lindquist@sofi.su.se

‡ University of Amsterdam and TI, J.Sol@uva.nl

§ Vrije Universiteit Amsterdam, Copenhagen Business School, CEPR, IZA, and TI, mvp.si@cbs.dk

A Online Appendix: Empirical Models

This Online Appendix provides a more formal description of our empirical approach. We first describe sibling correlations and their relationship with intergenerational correlations. We then adapt the method to compute neighborhood correlations and show how to account for observable parental characteristics to obtain estimates of their contribution to the sibling correlation. We also describe our peer effects model as well as our approach. Finally, we show how different types of sibling pairs can be used to construct an estimate of the contribution of shared genes to sibling correlations.

A.1 Sibling Correlations

Business ownership, E_{if}, for sibling i from family f can be modeled as:

$$E_{if} = X'_{if}\beta + \epsilon_{if}, \tag{A.1}$$

where X'_{if} includes individuals' birth year and a gender dummy for individual i from family f. The residual term, ϵ_{if}, is an individual-specific component representing a person's position in the overall distribution of business ownership, whose population variance is given by σ^2_ϵ. Following Solon (1999), the individual variance component, ϵ_{if}, is assumed to be comprised of two linearly additive and independent variance components:

$$\epsilon_{if} = a_f + b_{if}. \tag{A.2}$$

The first part, a_f, is a permanent component shared by all siblings in family f. This is what makes siblings similar. The second component, b_{if}, is the permanent component unique to sibling i in family f. The variance of ϵ_{if} can be expressed as the sum of the stationary population variances of the permanent family and individual components:

$$\sigma^2_\epsilon = \sigma^2_a + \sigma^2_b. \tag{A.3}$$

The share of the variance in an individual's long-run propensity to choose business ownership over wage employment that can be attributed to family background effects is:

$$\rho = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_b} \equiv \text{corr}(\epsilon_{if}, \epsilon_{i'f}). \tag{A.4}$$

This share coincides with the correlation in business ownership of randomly drawn pairs of siblings, which is why ρ is called a sibling correlation. This sibling correlation can be thought of as an omnibus measure of the importance of family and community effects. It includes family-wide influences that are shared by siblings, such as parental entrepreneurship, parental income, parental aspirations, cultural inheritance, genes, etc. However, it also includes shared influences that are not directly experienced in the home, such as school, church, and neighborhood effects. Genetic traits not shared by siblings, differential treatment of siblings, time-dependent changes in neighborhoods, schools, etc., are captured by the individual component b_{if}. If non-shared factors are relatively more important than shared factors for determining business ownership, the variance of family effects will be small relative to the variance of individual effects and the sibling correlation will be low; in other words, the more important the effects of factors that siblings share are, the larger the sibling correlation will be.[1]

[1] The existence of non-shared family factors, such as differential treatment by birth order, gender, or their interaction with family size, implies that sibling correlations should be viewed as lower bounds on the importance of family background and neighborhood effects. Björklund and Jäntti (2012) discuss this issue and examine the size of the advantage of first born children over their younger siblings in cognitive and non-cognitive skills, height, schooling, and earnings; they find only minor effects. Nonetheless, birth order effects could be important

An estimate of the sibling correlation in entrepreneurship entry, ρ, can be constructed using estimates of the between-family variation, σ_a^2, and the individual (within-family) variation, σ_b^2. These can be obtained by estimating the following latent linear response model:

$$E_{if}^* = \mathbf{X'}_{if}\beta + a_f + b_{if}, \tag{A.5}$$

where we only observe $E_{if} = I(E_{if}^* > 0)$ (i.e. the dependent variable is dichotomous). We estimate equation (5) using Stata's *xtlogit* command under the assumption that the random effect a_f is a realization from a normal distribution with mean zero and constant variance, while the individual variance component, b_{if}, is drawn from the logistic distribution with mean zero and variance $\pi^2/3$. Stata's *xtlogit* command reports ρ (along with a 95% confidence interval) as part of its standard output. For the continuous intensive margin outcomes, we estimate a similar model using Stata's *mixed* command under the assumption that the two random components are independent realizations from a multivariate normal distribution with mean zero and constant variance. The variance components are estimated using restricted maximum likelihood. These models are estimated only conditional on entry into self-employment.

A.2 Sibling Correlations and Intergenerational Correlations

Solon's (1999) derivation of the sibling correlation nicely demonstrates the analytical relationship between the intergenerational (e.g. parent-offspring) correlation, which we will call γ, and the sibling correlation, ρ. Let the permanent family component, a_f, be defined as the sum of parental business ownership (times γ), $\gamma\epsilon_f$, and a set of other parental factors orthogonal to ϵ_f, z_f. We then have:

$$a_f = \gamma\epsilon_f + z_f. \tag{A.6}$$

Taking the variance of both sides of equation (A.6) and dividing through by $\sigma_{\epsilon_{if}}^2$ gives us:

$$\frac{\sigma_{a_f}^2}{\sigma_{\epsilon_{if}}^2} = \rho = \frac{\gamma^2\sigma_{\epsilon_f}^2}{\sigma_{\epsilon_{if}}^2} + \frac{\sigma_{z_f}^2}{\sigma_{\epsilon_{if}}^2}. \tag{A.7}$$

If $\sigma_{\epsilon_f}^2 \approxeq \sigma_{\epsilon_{if}}^2$, then we obtain the following relationship:

$$\rho = \gamma^2 + \frac{\sigma_{z_f}^2}{\sigma_{\epsilon_{if}}^2}. \tag{A.8}$$

The sibling correlation equals the intergenerational correlation in business ownership squared plus all parental factors uncorrelated with parental business ownership. In fact, we show that the total effect of the latter dwarfs the importance of parental entrepreneurship when accounting for sibling correlations. Thus, focusing attention solely on intergenerational correlations results in a narrow measure of the overall importance of family for entrepreneurship.

A.3 Neighborhood Correlations

To calculate the contribution of neighborhoods to sibling correlations, we redefine the cluster as a neighborhood rather than a family. An estimate of the neighborhood correlation, ρ_n, can be constructed by using estimates of the between-neighborhood variation, σ_n^2, and the individual (within-neighborhood) variation, σ_b^2, which can be obtained by estimating the following latent

for self-employment (Black et al., 2018). Mishkin (2017) finds that in the U.S. the father-daughter association in self-employment is 80% lower if a brother is present. Vladasel (2018) studies birth order, family size, and sibling sex composition effects in entrepreneurship in Sweden and concludes that they are quantitatively unimportant. In particular, he finds that the mechanism described by Mishkin (2017) only applies to paternal unincorporated business ownership, and is much weaker in Sweden. Overall, sibling differences are unlikely to affect our estimates.

linear response model for our extensive margin outcomes:

$$E^*_{in} = \mathbf{X'}_{in}\beta + c_n + b_{in}, \tag{A.9}$$

where c_n is a permanent community factor and we only observe $E_{in} = I(E^*_{in} > 0)$. An equivalent model can be estimated for our continuous intensive margin outcomes. The main difference from previously estimated sibling correlations is that we also include a set of parental characteristics in $\mathbf{X'}_{in}$ to correct for parental sorting into neighborhoods (Solon et al., 2000). Correcting for sorting provides a tighter upper bound on neighborhood effects on business ownership outcomes. With these new variance components, the neighborhood correlation is calculated as:

$$\rho_n = \frac{\sigma^2_c}{\sigma^2_c + \sigma^2_b}. \tag{A.10}$$

As a last step, we take the ratio of neighborhood correlation to sibling correlation, ρ_n/ρ, as our metric for the contribution of neighborhood to the total importance of family and community background. In our empirical analysis, we define a neighborhood as the parish the individual resided in at age 15, the smallest geographical unit we observe in our data. Other neighborhood definitions, such as schools or statistical metropolitan areas are unlikely to induce large changes in these correlations (Raaum et al., 2006; Lindahl, 2011). Moreover, in our data, correlations estimated for wider definitions – municipalities and counties – are lower than parish correlations, suggesting we are accurately capturing a tightly-knit community that provides the individual's social context.

A.4 Accounting for Parental Characteristics

To examine which parental characteristics are mainly responsible for generating sibling similarities in business ownership, we include a set of family-wide variables suggested by the literature, either one at a time or simultaneously, as control variables in our regressions. Since parental characteristics are shared by children from the same family, we expect the family-level variance component to decrease. As a consequence the sibling correlation also decreases. The degree to which any particular control variable lowers the sibling correlation after being included provides a metric for judging its importance in explaining sibling similarities (Mazumder, 2008; Björklund et al., 2010) and the upper bound on its explanatory power, but does not allow for a strictly causal interpretation. Nonetheless, this exercise provides clues for the mechanisms that are relatively more important for siblings' entrepreneurship outcomes.

Consider the inclusion of mothers' and fathers' entrepreneurship in $\mathbf{X'}_{if}$. These two additional variables should reduce the residual variation in the outcome variable and produce a lower estimate of the between-family variation, σ^{2*}_a, than the estimate produced without the added controls. We can interpret the difference between these two estimates, $\sigma^2_a - \sigma^{2*}_a$, as an upper bound on the amount of the variance in the family component that can be explained by parental entrepreneurship. It is viewed as an upper bound since it includes other factors affecting children's entrepreneurship that are correlated with parental entrepreneurship (for instance, education, occupation, or residence). This exercise also produces a new sibling correlation, ρ^*. From what we know about the relationship between parents' and children's entrepreneurship, we expect this new sibling correlation to be substantially lower.

One could be concerned with a mechanical decrease in the sibling correlation as controls are added, similar to the mechanical increase in R^2. To ensure this is not the case, we generated a set of (20 or 100) noisy random variables at both the individual level and the family level, and included them as controls. Appendix Table B.6 shows that the change is not mechanical: these random variables increase (decrease) the sibling correlations by at most 1.59% (0.72%), and often explain nothing at all.

A.5 Peer Effects Model

We provide a formal exposition of the correlated random effects model suggested by Altonji et al. (2017) that we adopt for the purpose of estimating sibling peer effects in entrepreneurship. We begin by estimating the raw association between sibling i's unincorporated or incorporated business ownership at time t, $S_t^{i'}$, and sibling i''s business ownership at time $t-1$, $S_{t-1}^{i'}$:

$$S_t^i = \beta_0 + \beta_1 S_{t-1}^{i'} + u_t^2, \tag{A.11}$$

where the family subscript f is suppressed. We then add the set of controls used in the accounting exercise, X^f, and age dummies age_t^i for the focal sibling i:

$$S_t^i = \beta_0 + \beta_1 S_{t-1}^{i'} + X^f + age_t^i + \epsilon_t^2. \tag{A.12}$$

We estimate equations (A.11) and (A.12) (corresponding to columns (1), (2), (6) and (7) in Tables B.11 to B.15) by using the panel structure of our data, limiting the sample to families with two children.[2] We later split the sample into pairs of males, females, and mixed gender pairs, where the younger sibling is male or female. We use logistic regressions in order to maintain consistency with previous estimation techniques, and we report both odds ratios and (approximated) sibling correlations, as explained in the notes to Table B.11.

Part of the effect of sibling i''s business ownership on sibling i's business ownership estimated in equation (A.12), however, may be due to correlated random family effects, rather than direct peer effects. Altonji et al. (2017) suggest the use of a correlated random effects regression to isolate the direct sibling effect, achieving causal inference by assuming one-directional causation (whereas our study does not attempt to directly target causality); they control for the sum of sibling i''s business ownership at time $t-1$ and $t+1$ to net out the unobservable family component. We can then write:

$$S_t^i = \beta_0 + \beta_1 (S_{t-1}^{i'} + S_{t+1}^{i'}) + \lambda_0 S_{t-1}^{i'} + X^f + age_t^i + age_t^{i'} + \epsilon_t^2, \tag{A.13}$$

where the direct (lagged) sibling effect is captured by λ_0.[3] Similarly, we can also include a direct contemporaneous sibling influence by including sibling i''s business ownership at time t, $S_t^{i'}$, in conjunction with an expanded control for correlated random effects:

$$S_t^i = \beta_0 + \beta_1 (S_{t-1}^{i'} + S_t^{i'} + S_{t+1}^{i'}) + \lambda_0 S_{t-1}^{i'} + \lambda_1 S_t^{i'} + X^f + age_t^i + age_t^{i'} + \epsilon_t^2, \tag{A.14}$$

where λ_1 is the estimate of the 'contemporaneous' effect. This estimate should not be interpreted as a true contemporaneous effect, but rather as a transitory and common shock to both siblings in the same family. Hence, we do not sum the lagged and contemporaneous sibling effect when analyzing the contribution of peers to the sibling correlation (in contrast to Eriksson et al., 2016, for instance). Results for equations (A.13) and (A.14) are given in columns (3), (4), (8) and (9) of Tables B.11-B.15 below, while columns (5) and (10) present results from a variation of equation (A.11), where the lagged sibling effect is replaced by the contemporaneous one (this equation being necessary for calibration purposes).

As an example of how these tables should be interpreted, Table B.11 shows the results of our sibling peer effects exercise on the sub-sample of sibling pairs, with panel A referring to unincorporated business ownership and panel B to incorporation; in columns (1) to (5) sibling i is the younger one, whereas in columns (6)-(10), sibling i is the older one. The results suggest

[2] Sibling correlations for this sample are reported in row (7) of Appendix Table B.2 and closely match those reported in Table 3 in the paper.

[3] A detailed description of the assumptions and mechanics of this model is provided in Altonji et al. (2017). Importantly, they assume that only older siblings can influence the younger ones (and not the other way around), and that parental treatment of younger siblings does not change upon observing the behavior of older siblings.

a positive and significant (at 10 percent) impact of the younger sibling's incorporation status at time $t-1$ on the older sibling's incorporation status at time t, as shown in column (9) of Table B.11. This translates into a sibling correlation $\rho = 0.009$ as given by the lagged sibling effect, representing 2.85 percent of the baseline sibling correlation, as shown in column (4) of Table B.10.[4]

A.6 Genetic Effects Model

We begin by positing an additive model of genetic and environmental influences (see, e.g., Björklund et al., 2005). Business ownership, E, is due to a genetic factor, G, a shared environmental influence, S, and an unshared, idiosyncratic environmental influence, U:

$$E = gG + sS + uU, \tag{A.15}$$

where g, s, and u are model parameters representing the relative influence of each of these three factors. Given this model, our sibling correlation, $\rho = corr(E, E')$ is equal to $g^2 + s^2$, i.e. the share of the total variation in our entrepreneurial outcomes that can be attributed to shared environmental and genetic factors.

In our data, we have information on four different sibling types with different degrees of genetic relatedness: twins, full siblings, half siblings, and adopted siblings. Assuming that twins share on average 75% of their genes – since we pool monozygotic (MZ) and dizygotic (DZ) twins –, the twin correlation, ρ_{twin}, equals $0.75 * g^2 + s^2$. We assume that the correct models for full, half and adopted siblings are $\rho_{full} = 0.5 * g^2 + s^2$, $\rho_{half} = 0.25 * g^2 + s^2$, and $\rho_{adopted} = s^2$, respectively. We also assume that the correlation in shared environmental factors, $corr(S, S')$, is equal to one for all sibling types.[5]

We use a method of moments estimator to produce estimates of $\hat{g}^2$ and $\hat{s}^2$. This estimator chooses parameter estimates in order to match the four different sibling correlations observed in our data as best as possible (Björklund et al., 2005). We use weights to increase the precision of our estimates, where the weights are simply the sample size, N, that each sibling correlation is based on. This means that our estimator will work much harder to fit the model to the observed sibling correlation for full siblings than it will for the other three types of siblings.

In Table B.4, column (1) indicates sample size and column (2) shows the sibling correlation for a given sibling pair type. Column (3) then produces the estimated sibling correlation from the GMM estimation, split into a genetic component (g^2) in column (4) and a shared environment component (s^2) in column (5). Finally, column (6) displays the contribution of shared genes to the sibling correlation for each sibling type (g^2/ρ), while column (7) displays the weighted average of these contributions as our preferred estimate for the contribution of shared genes to sibling correlations in the full sample.

[4] While we argue that contemporaneous effects should not be included in the contribution of peer effects to the sibling correlation, they provide useful information. First, they are rarely significant for unincorporated self-employment, suggesting a limited role for common transitory shocks for this type of business. Second, they are always significant and positive for incorporation, and explain around 20% of the sibling correlation for all sibling types together (around 25% for brothers, 12.5% for sisters, and between 13% and 23% for mixed gender siblings). This result is consistent both with large common transitory shocks and with the possibility that siblings co-found businesses or take over the family firm. This suggests a rather loose upper bound on the importance of inheritance for our sibling correlations in the region of 25% for brothers, and smaller for other types of siblings. However, in this sample, siblings are observed as incorporated business owners at the same time and in the same industry in at most 3% of years.

[5] This exercise is designed to follow Björklund et al. (2005) as closely as possible. They, however, have data on nine different sibling types including some who are reared apart (i.e. whose environments are not perfectly correlated). This allows them to loosen and test some of the strict assumptions that we make in our simple additive model, producing lower estimates of genetic effects. Therefore, our results place an upper bound on the contribution of shared genes to sibling correlations.

B Online Appendix: Additional Figures and Tables

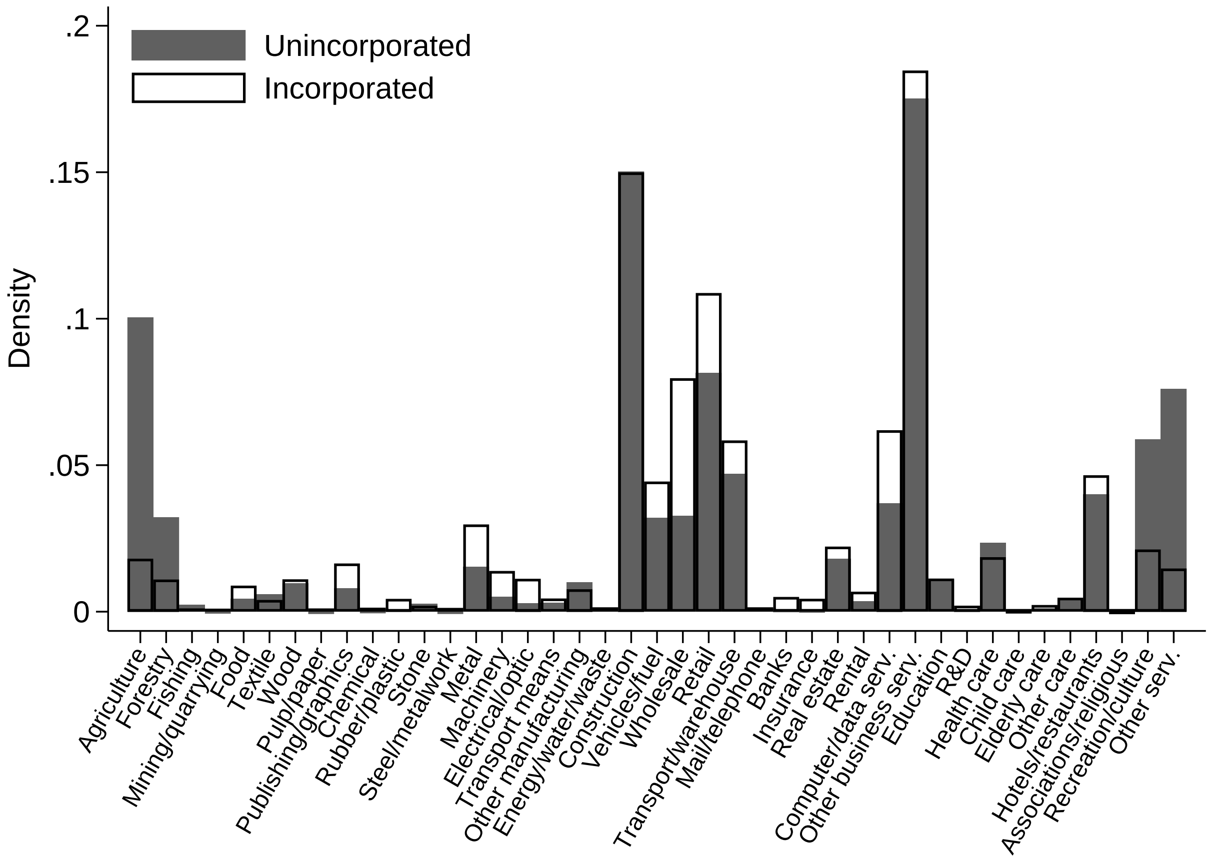

Figure B.1: Individual modal industries (1993-2010), by type of business.

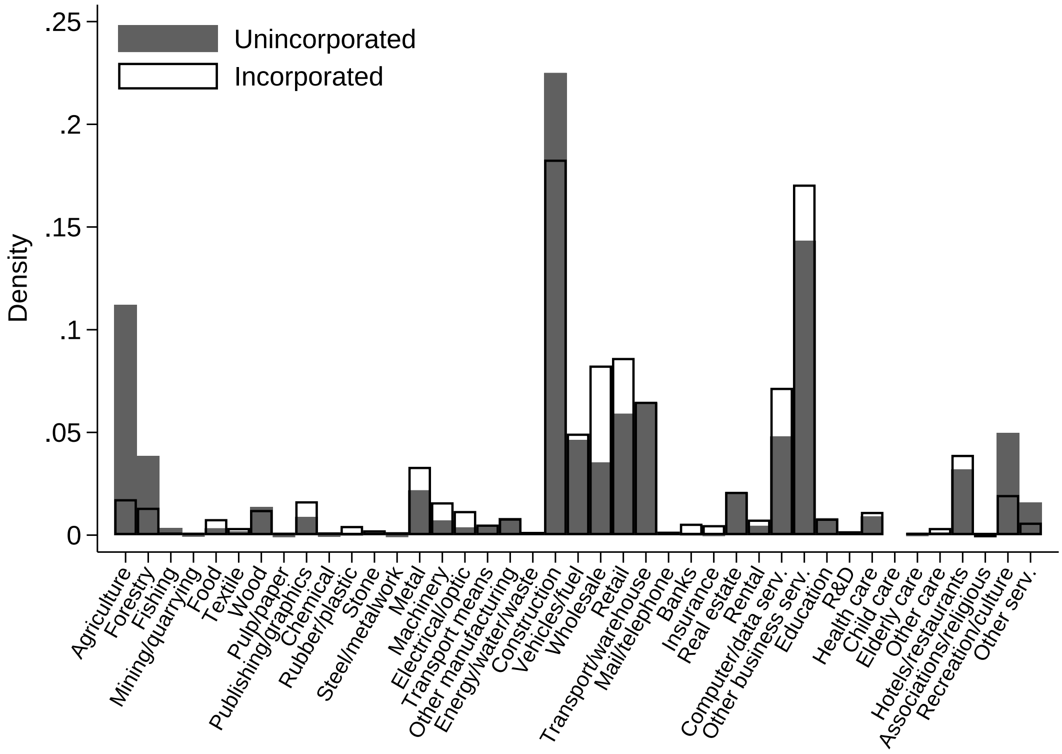

Figure B.2: Men's modal industries (1993-2010), by type of business.

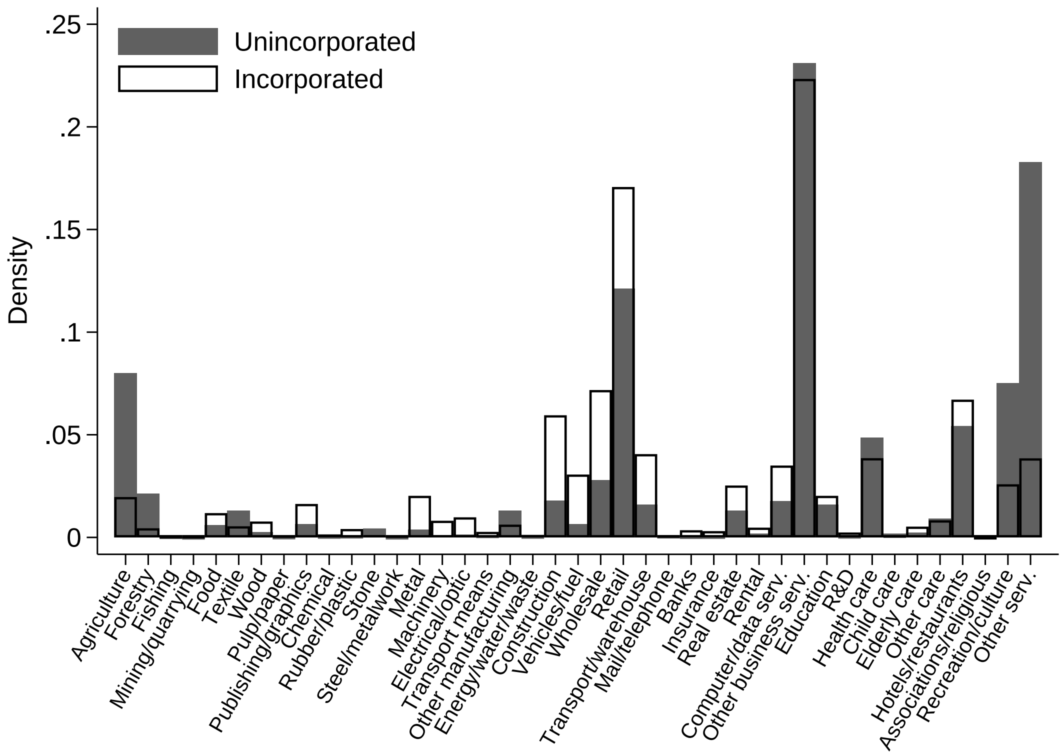

Figure B.3: Women's modal industries (1993-2010), by type of business.

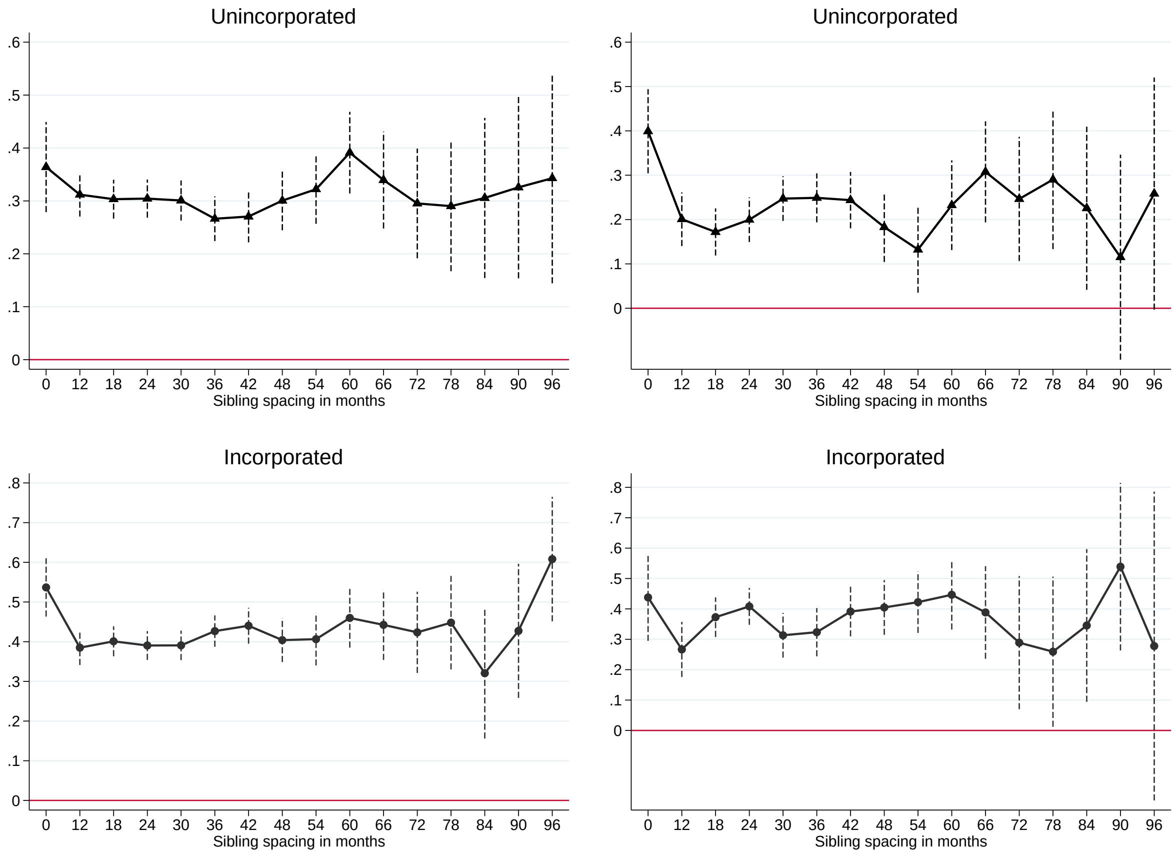

Figure B.4: Twin and sibling correlations in entrepreneurship by sibling spacing and gender (left – male; right – female), with 95% confidence intervals.

Table B.1: **Number of Families with *N* Children**

N Children	No. of Families	%	No. of Individuals	%
1	227,860	52.88	227,860	32.73
2	152,050	35.28	304,100	43.68
3	41.818	9.70	125,454	18.02
4	7,592	1.76	30,368	4.36
5	1,312	0.30	6,560	0.94
6	243	0.06	1,458	0.21
7	49	0.01	343	0.05
8	11	0.00	88	0.01
Total	430,935	100.00	696,231	100.00

All children of the same mother are defined as belonging to the same family.

Table B.2: **Sensitivity Analyses**

	Entry		Years	
All siblings	Unincorporated (1)	Incorporated (2)	Unincorporated (3)	Incorporated (4)
(1) Excl. singletons	0.212	0.342	0.215	0.387
	(0.004)	(0.007)	(0.009)	(0.009)
(2) Incl. 1985-1992	0.232		0.214	
	(0.004)		(0.008)	
(3) Outcomes, ages 25-40	0.222	0.369	0.222	0.403
	(0.005)	(0.007)	(0.011)	(0.014)
(4) Father	0.204	0.331	0.214	0.386
	(0.004)	(0.005)	(0.009)	(0.009)
(5) Excl. adoptive fathers	0.206	0.331	0.216	0.387
	(0.004)	(0.005)	(0.009)	(0.009)
(6) Excl. adoptive mothers	0.212	0.342	0.215	0.387
	(0.004)	(0.005)	(0.009)	(0.009)
(7) Non-twin pairs	0.214	0.336	0.210	0.408
	(0.006)	(0.007)	(0.014)	(0.012)
(8) Non-twins, 12-24 months	0.208	0.298	0.230	0.430
	(0.012)	(0.015)	(0.028)	(0.027)
(9) Complete families	0.200	0.341	0.216	0.382
	(0.006)	(0.007)	(0.014)	(0.012)
(10) Complete parental data	0.208	0.341	0.213	0.394
	(0.004)	(0.005)	(0.010)	(0.009)
(11) Birth year dummies	0.211	0.340	0.214	0.388
	(0.004)	(0.005)	(0.009)	(0.009)
(12) Parents $\leq$ 65yo in 1993	0.210	0.341	0.202	0.379
	(0.005)	(0.005)	(0.010)	(0.010)
(13) Parents in data in 1993	0.212	0.341	0.201	0.392
	(0.005)	(0.006)	(0.011)	(0.010)
(14) Excl. farmer parents	0.207	0.340	0.182	0.378
	(0.005)	(0.005)	(0.010)	(0.010)
(15) 'Fake' families	0.002	0.003	0.004	0.005
	(0.002)	(0.005)	(0.006)	(0.007)

Standard errors in parentheses. Row (1) excludes singletons; row (2) includes data on unincorporated business ownership, 1985-1992; row (3) measures the outcomes only between ages 25 and 40; row (4) defines the family through the father; row (5) omits families with an adoptive father; row (6) omits families with an adoptive mother; row (7) restricts the analysis to families with two children; row (8) restricts it further to closely spaced non-twin pairs (born 12 to 24 months apart); row (9) restricts the analysis to families completely captured in our sample; row (10) restricts the analysis to observations for which all parental characteristics are observed; row (11) includes individual and parental birth year dummies; row (12) retains families with parents below age 65 in 1993; row (13) retails families with parents who have not left the data set and are below age 65 in 1993; row (14) drops families where one of the parents is a farmer; finally, row (15) is a placebo test, where the family cluster structure is replicated and individuals randomly allocated to families, with 100 bootstrap replications.

Table B.3: **Sibling Correlations in Entrepreneurship, by Family Type**

	Entry		Years		Income	
	Unincorporated	Incorporated	Unincorporated	Incorporated	Unincorporated	Incorporated
	(1)	(2)	(3)	(4)	(5)	(6)
A. Male only sibships						
ρ	0.269	0.399	0.318	0.463	0.291	0.415
	(0.012)	(0.012)	(0.023)	(0.018)	(0.026)	(0.021)
Individuals	69,340	69,340	10,597	8,466	10,597	8,466
Families	43,983	43,983	9,446	7,363	9,446	7,363
B. Female only sibships						
ρ	0.206	0.351	0.167	0.358	0.113	0.345
	(0.017)	(0.022)	(0.045)	(0.046)	(0.045)	(0.046)
Individuals	62,296	62,296	6,185	2,923	6,185	2,923
Families	40,304	40,304	5,783	2,747	5,783	2,747
C. Mixed gender sibships						
ρ	0.173	0.311	0.175	0.321	0.202	0.304
	(0.008)	(0.009)	(0.018)	(0.018)	(0.018)	(0.021)
Individuals	138,654	138,654	17,729	12,423	17,729	12,423
Families	57,416	57,416	15,258	10,831	15,258	10,831

Standard errors in parentheses. The sample is restricted to families captured completely in our sample in order to correctly define family type (by child gender: male only, female only, mixed).

Table B.4: **Shared Genes and Sibling Correlations at the Extensive Margin**

Sibling type	N (1)	$\hat{\rho}$ (2)	Model ρ (3)	g^2 (4)	s^2 (5)	g^2/ρ (6)	Average g^2/ρ (7)
A. Brothers, unincorporated				0.293	0.153		**47%**
				(0.134)	(0.073)		
MZ+DZ twins	4,416	0.363	0.373	0.220	0.153	59%	
		(0.043)					
Full siblings	37,772	0.306	0.300	0.147	0.153	49%	
		(0.014)					
Half siblings	5,036	0.157	0.227	0.073	0.153	32%	
		(0.041)					
Adopted siblings	1,010	0.304	0.153	0.000	0.153	0%	
		(0.087)					
B. Sisters, unincorporated				0.340	0.058		**72%**
				(0.147)	(0.069)		
MZ+DZ twins	4,438	0.399	0.312	0.255	0.058	82%	
		(0.048)					
Full siblings	33,288	0.206	0.227	0.170	0.058	75%	
		(0.020)					
Half siblings	4,746	0.201	0.143	0.085	0.058	60%	
		(0.056)					
Adopted siblings	1,126	0.105	0.058	0.000	0.058	0%	
		(0.118)					
C. Brothers, incorporated				0.391	0.196		**48%**
				(0.064)	(0.030)		
MZ+DZ twins	4,416	0.528	0.490	0.294	0.196	60%	
		(0.039)					
Full siblings	37,772	0.384	0.392	0.196	0.196	50%	
		(0.015)					
Half siblings	5,036	0.317	0.294	0.098	0.196	33%	
		(0.049)					
Adopted siblings	1,010	0.223	0.196	0.000	0.196	0%	
		(0.107)					
D. Sisters, incorporated				0.413	0.137		**58%**
				(0.195)	(0.107)		
MZ+DZ twins	4,438	0.429	0.447	0.309	0.137	69%	
		(0.072)					
Full siblings	33,288	0.355	0.344	0.206	0.137	60%	
		(0.026)					
Half siblings	4,746	0.130	0.240	0.103	0.137	43%	
		(0.122)					
Adopted siblings	1,126	0.335	0.137	0.000	0.137	0%	
		(0.142)					

Standard errors in parentheses. GMM estimation of genetic effects using different types of siblings (Björklund et al., 2005). The last column shows the weighted average contribution of shared genes.

Table B.5: **Accounting Exercise: Parental Characteristics**

	Unincorporated		Incorporated	
	Brothers	Sisters	Brothers	Sisters
	(1)	(2)	(3)	(4)
A. Parental education				
Mother's education level:				
9 years	1.074***	1.190***	1.344***	1.278***
11 years	1.077***	1.098***	1.241***	1.118***
12 years	1.195***	1.410***	1.591***	1.507***
14 years	1.545***	1.283***	1.536***	1.349***
15.5 years	1.205***	1.423***	1.555***	1.409***
19 years	1.579***	1.548***	1.658***	1.404***
Missing	1.470***	1.141***	1.088	0.965
Father's education level:				
9 years	1.087***	1.176***	1.203***	1.374***
11 years	1.050***	1.115***	1.044**	1.127***
12 years	0.865***	1.200***	1.284***	1.429***
14 years	0.842***	1.186***	1.121***	1.417***
15.5 years	0.838***	1.390***	1.143***	1.483***
19 years	0.834***	1.490***	1.133**	1.606***
Missing	1.122***	1.218***	0.754***	0.980
$\rho*$	0.289	0.205	0.399	0.341
	(0.007)	(0.010)	(0.007)	(0.013)
	0.86%	3.66%	1.13%	2.98%
B. Parental income				
Pct. 10-20	0.890***	0.996	1.180***	1.190***
Pct. 20-30	0.752***	0.928**	1.322***	1.191***
Pct. 30-40	0.700***	0.929**	1.389***	1.399***
Pct. 40-50	0.635***	0.930**	1.488***	1.340***
Pct. 50-60	0.643***	0.895***	1.637***	1.441***
Pct. 60-70	0.632***	0.915***	1.809***	1.548***
Pct. 70-80	0.647***	1.024	2.182***	1.880***
Pct. 80-90	0.660***	1.180***	2.718***	2.388***
Pct. 90-95	0.662***	1.409***	3.219***	3.022***
Pct. 95-99	0.689***	1.663***	4.049***	4.232***
Pct. 99-100	0.770***	2.128***	6.354***	7.219***
Missing	0.774***	0.992	1.033***	1.239***
$\rho*$	0.288	0.206	0.382	0.321
	(0.007)	(0.010)	(0.007)	(0.013)
	1.47%	3.14%	5.41%	8.67%

continued

Table B.5 (cont'd): **Accounting Exercise: Parental Characteristics**

	Unincorporated		Incorporated	
	Brothers	Sisters	Brothers	Sisters
	(1)	(2)	(3)	(4)
C. Parental entrepreneurship				
Mother unincorporated	1.643***	1.589***	1.286***	1.282***
Father unincorporated	1.963***	1.356***	1.342***	1.196***
Mother incorporated	0.911***	1.294***	2.837***	3.507***
Father incorporated	0.933***	1.231***	3.854***	2.032***
Missing (father)	1.467***	1.118**	0.740***	0.815**
$\rho*$	0.267	0.199	0.348	0.309
	(0.007)	(0.010)	(0.008)	(0.014)
	8.66%	6.37%	13.75%	12.15%
D. Other family characteristics				
Mother immigrant	1.088***	1.010	0.845***	0.913**
Father immigrant	1.153***	1.044*	0.970***	0.945
Missing (father)	1.111**	0.968	0.702***	0.804**
Family structure:				
Single mother	1.046***	0.990	0.746***	0.763***
Single father	1.136***	1.102***	0.919***	0.882**
Mother, new husband	1.052*	1.044	0.784***	0.866***
Father, new wife	1.127***	1.031	0.587***	0.672***
Missing	0.940	0.865**	0.447***	0.536***
$\rho*$	0.290	0.213	0.399	0.348
	(0.007)	(0.010)	(0.007)	(0.013)
	0.56%	0.05%	1.18%	1.06%
E. All family characteristics				
$\rho*$	0.257	0.190	0.334	0.287
	(0.007)	(0.010)	(0.008)	(0.014)
	11.63%	10.75%	17.23%	18.20%
Individuals	356,847	339,384	356,847	339,384
Families	278,107	267,894	278,107	267,894

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses. Percentages indicate the contribution of parental characteristics to the corresponding sibling correlations in columns (1) and (2) of Table 3 in the paper. For parental education, the reference category is 7 years; for parental income, the reference category is the bottom decile of parental income distribution; for family structure, the reference category is the intact family. The models include dummies for missing parental education, income, business ownership, immigrant status, and family structure, when assessing the explanatory effect of that particular family variable. In Panel C, we also include dummy variables for the mother and father leaving the sample before 1985 (when data on unincorporated business ownership becomes available) and before 1993 (when data on incorporated business ownership becomes available).

Table B.6: **Accounting Exercise: Effect of Noisy Random Variables**

	Entry		Years		Income	
	(1) Uninc.	(2) Inc.	(3) Uninc.	(4) Inc.	(5) Uninc.	(6) Inc.
A. Individual level: 20 random variables in (0,1)						
Brothers	0.292	0.404	0.309	0.448	0.297	0.416
	(0.007)	(0.007)	(0.013)	(0.011)	(0.014)	(0.014)
	0.01%	0.02%	0.05%	-0.21%	-0.11	-0.11%
Sisters	0.213	0.352	0.170	0.380	0.158	0.360
	(0.010)	(0.013)	(0.026)	(0.030)	(0.026)	(0.033)
	0.03%	-0.06%	-0.55%	0.30%	-0.46%	0.04%
B. Individual level: 100 random variables in (0,1)						
Brothers	0.292	0.404	0.309	0.446	0.298	0.416
	(0.007)	(0.007)	(0.013)	(0.011)	(0.014)	(0.013)
	0.05%	-0.11%	0.04%	0.11%	-0.48%	-0.02%
Sisters	0.213	0.352	0.169	0.385	0.160	0.362
	(0.010)	(0.013)	(0.026)	(0.030)	(0.026)	(0.034)
	-0.12%	-0.08%	-0.16%	-1.16%	-1.59%	-0.53%
C. Family level: 20 random variables in (0,1)						
Brothers	0.292	0.403	0.309	0.447	0.296	0.416
	(0.007)	(0.007)	(0.013)	(0.011)	(0.014)	(0.013)
	0.06%	0.04%	0.09%	0.00%	0.03%	-0.01%
Sisters	0.213	0.351	0.168	0.381	0.157	0.361
	(0.010)	(0.013)	(0.026)	(0.029)	(0.026)	(0.033)
	0.09%	0.07%	0.17%	-0.06%	0.63%	-0.07%
D. Family level: 100 random variables in (0,1)						
Brothers	0.291	0.403	0.309	0.446	0.296	0.415
	(0.007)	(0.007)	(0.013)	(0.011)	(0.014)	(0.013)
	0.24%	0.09%	0.14%	0.09%	0.17%	0.10%
Sisters	0.212	0.350	0.170	0.378	0.157	0.361
	(0.010)	(0.013)	(0.026)	(0.030)	(0.026)	(0.033)
	0.53%	0.38%	-0.58%	0.72%	0.11%	-0.08%

Standard errors in parentheses. The percentages indicate the change in the sibling correlations in Table 3 once the noisy random variables are controlled for (a negative percentage change indicates an increase in the sibling correlation, while a positive sign indicates a decrease). In panels A and B, the noisy random variables are generated at the individual level, and the largest change in the sibling correlation is a 1.59% *increase*. In panels C and D, the noisy random variables are generated at the family level, such that they are the same for siblings, and should have a higher explanatory power than those generated at the individual level. Even so, they explain at most 0.72% of the sibling correlation. While the number of variables appears inconsequential when variables are generated at the individual level, when they are generated at the family level the decrease is slightly larger when 100 variables are added instead of 20. Our models in Online Appendix Table B.5 include far less than 100 variables, suggesting little cause for concern that the explanatory power is generated simply by random noise.

Table B.7: **Accounting Exercise, Extensive and Intensive Margin Outcomes**

		Entry		Years		Income	
		(1) Uninc.	(2) Inc.	(3) Uninc.	(4) Inc.	(5) Uninc.	(6) Inc.
A. Parental education	All	0.210	0.335	0.208	0.379	0.190	0.321
		(0.004)	(0.005)	(0.009)	(0.009)	(0.010)	(0.011)
		1.07%	1.84%	3.06%	1.91%	9.18%	7.17%
	Brothers	0.289	0.399	0.300	0.437	0.273	0.390
		(0.007)	(0.007)	(0.013)	(0.011)	(0.014)	(0.014)
		0.86%	1.13%	3.14%	2.07%	7.88%	6.12%
	Sisters	0.205	0.341	0.167	0.379	0.149	0.347
		(0.010)	(0.013)	(0.026)	(0.030)	(0.027)	(0.034)
		3.66%	2.98%	0.99%	0.47%	5.62%	3.67%
B. Parental income	All	0.209	0.315	0.204	0.386	0.169	0.325
		(0.004)	(0.005)	(0.009)	(0.009)	(0.010)	(0.011)
		1.46%	7.58%	4.80%	0.21%	19.00%	5.96%
	Brothers	0.288	0.382	0.294	0.446	0.249	0.397
		(0.007)	(0.007)	(0.013)	(0.011)	(0.015)	(0.014)
		1.47%	5.41%	4.99%	0.20%	15.87%	4.50%
	Sisters	0.206	0.321	0.168	0.378	0.139	0.340
		(0.010)	(0.013)	(0.026)	(0.030)	(0.027)	(0.034)
		3.14%	8.67%	0.51%	0.75%	12.01%	5.49%
C. Parental entrepreneurship	All	0.189	0.282	0.196	0.352	0.197	0.332
		(0.004)	(0.005)	(0.009)	(0.010)	(0.010)	(0.011)
		10.67%	17.26%	8.38%	8.81%	5.62%	4.10%
	Brothers	0.267	0.348	0.284	0.413	0.279	0.399
		(0.007)	(0.008)	(0.013)	(0.012)	(0.014)	(0.014)
		8.66%	13.75%	8.12%	7.47%	5.80%	4.03%
	Sisters	0.199	0.309	0.164	0.350	0.154	0.354
		(0.010)	(0.014)	(0.026)	(0.031)	(0.026)	(0.033)
		6.37%	12.15%	2.79%	8.07%	2.46%	1.65%
D. Other family traits	All	0.211	0.337	0.209	0.381	0.203	0.345
		(0.004)	(0.005)	(0.009)	(0.009)	(0.010)	(0.011)
		0.39%	1.34%	2.28%	1.46%	2.68%	0.23%
	Brothers	0.290	0.399	0.304	0.440	0.291	0.415
		(0.007)	(0.007)	(0.013)	(0.011)	(0.014)	(0.013)
		0.56%	1.18%	1.81%	1.38%	1.86%	0.23%
	Sisters	0.213	0.348	0.168	0.379	0.153	0.360
		(0.010)	(0.013)	(0.026)	(0.030)	(0.027)	(0.033)
		0.05%	1.06%	0.49%	0.63%	3.02%	0.09%
E. All family controls	All	0.183	0.266	0.184	0.344	0.158	0.228
		(0.004)	(0.005)	(0.009)	(0.010)	(0.010)	(0.012)
		13.48%	22.05%	13.90%	11.08%	24.34%	16.62%
	Brothers	0.258	0.334	0.268	0.403	0.234	0.359
		(0.007)	(0.008)	(0.013)	(0.012)	(0.015)	(0.015)
		11.63%	17.23%	13.27%	9.83%	20.93%	13.66%
	Sisters	0.190	0.287	0.163	0.349	0.133	0.319
		(0.010)	(0.014)	(0.026)	(0.031)	(0.027)	(0.036)
		10.75%	18.20%	3.40%	8.42%	15.64%	11.36%

Standard errors in parentheses. The percentages indicate the contribution of parental characteristics to the corresponding sibling correlations in Table 3 in the paper.

Table B.8: **Neighborhood Correlations in Entrepreneurship**

	Unincorporated		Incorporated	
	Brothers	Sisters	Brothers	Sisters
	(1)	(2)	(3)	(4)
A. No controls				
	0.038	0.018	0.020	0.022
	(0.002)	(0.001)	(0.001)	(0.002)
	13.11%	8.40%	5.06%	6.24%
N	352,145	335,173	352,145	335,147
B. Parental controls (excl. entrepreneurship)				
	0.036	0.015	0.016	0.013
	(0.002)	(0.001)	(0.001)	(0.002)
	12.46%	7.04%	3.98%	3.75%
N	352,145	335,173	352,145	335,147
C. Parental controls (incl. entrepreneurship)				
	0.023	0.011	0.011	0.010
	(0.001)	(0.001)	(0.001)	(0.001)
	7.99%	5.21%	2.60%	2.75%
N	352,145	335,173	352,145	335,147
D. All controls (incl. share of entrepreneurs in neighborhood)				
	0.018	0.008	0.006	0.006
	(0.001)	(0.001)	(0.001)	(0.001)
	6.10%	3.81%	1.58%	1.62%
N	347,702	330,799	347,702	330,799
Odds ratios:				
% Other parents uninc.	3.581***	2.023***	1.315***	1.389***
% Other parents inc.	1.953**	10.508***	52.689***	50.824***

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses. Percentages indicate the contribution of neighborhood correlations to the corresponding sibling correlations in columns (1) and (2) of Table 3 in the paper. Panel D displays the odds ratios of the share of parents of individuals in the same parish who have been unincorporated or incorporated (leaving out the focal individual's parents).

Table B.9: **Neighborhood Correlations, Extensive and Intensive Margins**

	Entry		Years		Income	
	(1) Uninc.	(2) Inc.	(3) Uninc.	(4) Inc.	(5) Uninc.	(6) Inc.
A. No controls						
All	0.028	0.020	0.038	0.027	0.021	0.027
	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)
	13.03%	5.89%	17.90%	7.03%	9.94%	7.67%
Brothers	0.038	0.020	0.073	0.036	0.037	0.031
	(0.002)	(0.001)	(0.004)	(0.003)	(0.003)	(0.003)
	13.11%	5.06%	23.67%	7.96%	12.59%	7.36%
Sisters	0.018	0.022	0.002	0.013	0.007	0.020
	(0.001)	(0.002)	(0.001)	(0.004)	(0.002)	(0.004)
	8.40%	6.24%	1.16%	3.30%	4.14%	5.55%
B. Parental controls (excl. entrepreneurship)						
All	0.027	0.015	0.028	0.021	0.011	0.011
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
	12.76%	4.53%	13.27%	5.48%	5.13%	3.32%
Brothers	0.036	0.016	0.056	0.027	0.021	0.013
	(0.002)	(0.001)	(0.004)	(0.003)	(0.002)	(0.002)
	12.46%	3.98%	18.18%	5.96%	7.21%	3.07%
Sisters	0.015	0.013	0.002	0.011	0.002	0.008
	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.003)
	7.04%	3.75%	1.07%	2.98%	1.58%	2.16%
C. Parental controls (incl. entrepreneurship)						
All	0.017	0.010	0.017	0.016	0.007	0.009
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
	8.10%	2.99%	7.81%	4.18%	3.58%	2.48%
Brothers	0.023	0.011	0.034	0.020	0.014	0.009
	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)
	7.99%	2.60%	10.96%	4.52%	4.79%	2.22%
Sisters	0.011	0.010	0.002	0.011	0.002	0.007
	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.003)
	5.21%	2.75%	1.05%	2.76%	1.54%	1.98%
D. All controls (incl. share of entrepreneurs in neighborhood)						
All	0.013	0.006	0.012	0.014	0.006	0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	6.10%	1.84%	5.44%	3.69%	2.99%	1.88%
Brothers	0.018	0.006	0.022	0.018	0.011	0.007
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
	6.10%	1.58%	7.20%	3.93%	3.70%	1.59%
Sisters	0.018	0.006	0.002	0.009	0.002	0.006
	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.003)
	3.81%	1.62%	0.97%	2.42%	1.34%	1.55%

Standard errors in parentheses. The percentages indicate the contribution of neighborhood correlations to the corresponding sibling correlations in Table 3 in the paper.

Table B.10: **Upper Bounds on Peer Effects Contribution to Sibling Correlations**

	Effect on younger sibling		Effect on older sibling	
	(1)	(2)	(3)	(4)
A. Unincorporated				
All sibling types	5.11	5.05	-2.07	-3.71
Males	7.24**	7.82**	-1.19	-1.97
Females	10.29	9.78	4.04	-0.55
Mixed (younger brother)	-9.99	-8.36	-2.68	-3.70
Mixed (younger sister)	0.72	-1.84	-12.02	-14.52
B. Incorporated				
All sibling types	-2.78	-1.37	2.12	2.85*
Males	-0.61	0.86	0.20	1.45
Females	-8.93*	-7.12	1.61	1.92
Mixed (younger brother)	-5.86	-4.70	7.51	7.73
Mixed (younger sister)	-6.15	-4.36	7.76	7.33
Contemporaneous effect	No	Yes	No	Yes

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All numbers are in percentages, representing the share of the sibling correlation explained by the lagged entrepreneurship status of the older sibling, columns (1) and (2), and the younger sibling, columns (3) and (4), once controls are added and correlated random effects are accounted for. For the full set of results, see Online Appendix Tables B.11-B.15; the results in this table are based on columns (3), (4), (8) and (9) in those tables. Note that applying a Bonferroni correction for testing multiple hypotheses (i.e., given that we estimate 40 different models, requiring a p-value below 0.00125 = 0.05/40) would render all estimates insignificant.

Table B.11: **Peer Effects Exercise, All Sibling Types**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Unincorporated										
	Old on young [$\phi_{t-1} = 0.084, \phi_t = 0.082$]					Young on old [$\phi_{t-1} = 0.084, \phi_t = 0.082$]				
OR($S_{t-1}^{i'}$)	3.096***	2.510***	1.065	1.065		3.067***	2.487***	0.975	0.955	
ρ	0.208	0.166	0.011	0.011		0.208	0.166	-0.004	-0.008	
OR($S_t^{i'}$)				1.041	3.185***				0.955	3.185***
ρ				0.007	0.208				-0.007	0.208
B. Incorporated										
	Old on young [$\phi_{t-1} = 0.068, \phi_t = 0.066$]					Young on old [$\phi_{t-1} = 0.068, \phi_t = 0.066$]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S_{t-1}^{i'}$)	7.802***	4.584***	0.936	0.968		0.942***	4.583***	1.052	1.070*	
ρ	0.324	0.231	-0.009	-0.004		0.324	0.229	0.007	0.009	
OR($S_t^{i'}$)				1.650***	8.412***				1.632***	8.412***
ρ				0.068	0.324				0.067	0.324
Family background		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i'			Yes	Yes				Yes	Yes	
$S_{t-1}^{i'} + S_{t+1}^{i'}$			Yes					Yes		
$S_{t-1}^{i'} + S_t^{i'} + S_{t+1}^{i'}$				Yes					Yes	

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. In columns (1)-(5), sibling i is the younger sibling; in columns (6)-(10), sibling i is the older sibling. The odds ratios (OR) are estimated using logistic regressions. Family background variables are those used in the accounting exercise: parental education, income, immigration, ownership of an unincorporated or incorporated business, and the family structure variables. Sibling correlations in columns (1), (5), (6) and (10) are estimated using Stata's *xtlogit* command; those in columns (2)-(4) and (7)-(9) are approximated using the following formula (Bonett, 2007): $\rho \approx (OR^{\phi} + 1)(OR^{\phi} - 1)$, where ϕ is calibrated using the odds ratios and sibling correlations estimated in columns (1) and (6) for lagged effects (ϕ_{t-1}), and (5) and (10) for contemporaneous effects (ϕ_t).

Table B.12: **Peer Effects Exercise, Brothers**

	A. Unincorporated									
	Old on young [$\phi_{t-1} = 0.084, \phi_t = 0.082$]					Young on old [$\phi_{t-1} = 0.084, \phi_t = 0082$]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S_{t-1}^{i'}$)	4.530***	3.596***	1.131**	1.142**		4.487***	3.542***	0.980	0.967	
ρ	0.289	0.240	0.021	0.023		0.289	0.238	-0.003	-0.006	
OR($S_t^{i'}$)				1.178**	4.723***				1.029	4.723***
ρ				0.027	0.289				0.005	0.289
	B. Incorporated									
	Old on young [$\phi_{t-1} = 0.075, \phi_t = 0.073$]					Young on old [$\phi_{t-1} = 0.075, \phi_t = 0.73$]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S_{t-1}^{i'}$)	9.542***	5.754***	0.984	1.023		9.646***	5.730***	1.006	1.040	
ρ	0.405	0.302	-0.002	0.003		0.405	0.300	0.001	0.006	
OR($S_t^{i'}$)				1.945***	10.395***				1.821***	10.395***
ρ				0.102	0.405				0.091	0.405
Family background		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i'			Yes	Yes				Yes	Yes	
$S_{t-1}^{i'} + S_{t+1}^{i'}$			Yes					Yes		
$S_{t-1}^{i'} + S_t^{i'} + S_{t+1}^{i'}$				Yes					Yes	

See notes to Table B.11. The p-values of the significant odds ratios of $S_{t-1}^{i'}$ in columns (3) and (4) of panel A are 0.045 and 0.034, respectively.

Table B.13: **Peer Effects Exercise, Sisters**

	A. Unincorporated									
	Old on young [$\phi_{t-1} = 0.087, \phi_t = 0.085$]					Young on old [$\phi_{t-1} = 0.088, \phi_t = 0.085$]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S_{t-1}^{i'}$)	3.136***	2.668***	1.137	1.130		3.095***	2.629***	1.052	0.993	
ρ	0.221	0.187	0.023	0.022		0.221	0.186	0.009	-0.001	
OR($S_t^{i'}$)				1.120	3.259***				1.023	3.259***
ρ				0.019	0.221				0.004	0.221
	B. Incorporated									
	Old on young [$\phi_{t-1} = 0.062, \phi_t = 0.060$]					Young on old [$\phi_{t-1} = 0.061, \phi_t = 0.060$]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S_{t-1}^{i'}$)	11.005***	6.870***	0.776*	0.818		11.509***	7.107***	1.047	1.056	
ρ	0.345	0.269	-0.031	-0.025		0.345	0.268	0.006	0.007	
OR($S_t^{i'}$)				1.428*	11.990***				1.459*	11.990***
ρ				0.043	0.345				0.046	0.345
Family background		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i'			Yes	Yes				Yes	Yes	
$S_{t-1}^{i'} + S_{t+1}^{i'}$			Yes					Yes		
$S_{t-1}^{i'} + S_t^{i'} + S_{t+1}^{i'}$				Yes					Yes	

See notes to Table B.11.

Table B.14: **Peer Effects Exercise, Mixed (Younger Brother)**

	A. Unincorporated									
	Old on young [$\phi_{t-1} =, \phi_t =$]					Young on old [$\phi_{t-1} =, \phi_t =$]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S^{i'}_{t-1}$)	2.026***	1.772***	0.927	0.939		2.043***	1.784***	0.980	0.972	
ρ	0.133	0.107	-0.013	-0.011		0.133	0.107	-0.004	-0.005	
OR($S^{i'}_{t}$)				0.852	2.063***				0.914	2.062***
ρ				-0.027	0.133				-0.015	0.133
	B. Incorporated									
	Old on young [$\phi_{t-1} =, \phi_t = 8$]					Young on old [$\phi_{t-1} =, \phi_t =$]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S^{i'}_{t-1}$)	5.491***	3.051***	0.896	0.916		5.587***	3.121***	1.150	1.155	
ρ	0.189	0.120	-0.011	-0.009		0.189	0.121	0.014	0.015	
OR($S^{i'}_{t}$)				1.351**	5.792***				1.309**	5.792***
ρ				0.030	0.189				0.027	0.189
Family background		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i'			Yes	Yes				Yes	Yes	
$S^{i'}_{t-1} + S^{i'}_{t+1}$			Yes					Yes		
$S^{i'}_{t-1} + S^{i'}_{t} + S^{i'}_{t+1}$				Yes					Yes	

See notes to Table B.11.

Table B.15: **Peer Effects Exercise, Mixed (Younger Sister)**

	A. Unincorporated									
	Old on young $[\phi_{t-1} = 0.078, \phi_t = 0.078]$					Young on old $[\phi_{t-1} = 0.080, \phi_t = 0.078]$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S^{i'}_{t-1}$)	2.230***	1.950***	1.006	0.984		2.178***	1.922***	0.904	0.886	
ρ	0.133	0.110	0.001	-0002		0.133	0.111	-0.016	-0.019	
OR($S^{i'}_{t}$)				0.915	2.235***				0.793**	2.235***
ρ				-0.014	0.133				-0.035	0.133
	B. Incorporated									
	Old on young $[\phi_{t-1} = 0.054, \phi_t = 0.052]$					Young on old $[\phi_{t-1} = 0.053, \phi_t = 0.052]$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OR($S^{i'}_{t-1}$)	4.950***	3.040***	0.898	0.926		5.111***	3.126***	1.147	1.138	
ρ	0.189	0.128	-0.012	-0.008		0.189	0.128	0.015	0014	
OR($S^{i'}_{t}$)				1.273*	5.286***				1.518***	5.286***
ρ				0.025	0.189				0.044	0.189
Family background		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i		Yes	Yes	Yes			Yes	Yes	Yes	
Age dummies, i'			Yes	Yes				Yes	Yes	
$S^{i'}_{t-1} + S^{i'}_{t+1}$			Yes					Yes		
$S^{i'}_{t-1} + S^{i'}_{t} + S^{i'}_{t+1}$				Yes					Yes	

See notes to Table B.11.

References

Altonji, J. G., Cattan, S., and Ware, I. (2017). Identifying Sibling Influence on Teenage Substance Use. *Journal of Human Resources*, 52(1):1–47.

Björklund, A. and Jäntti, M. (2012). How Important is Family Background for Labor-Economic Outcomes? *Labour Economics*, 19(4):465–474.

Björklund, A., Jäntti, M., and Solon, G. (2005). Influences of Nature and Nurture on Earnings Variation: A Report on a Study of Various Sibling Types in Sweden. In Bowles, S., Gintis, H., and Osborne, M., editors, *Unequal Chances: Family Background and Economic Success*, pages 145–164. Russel Sage Foundation, New York.

Björklund, A., Lindahl, L., and Lindquist, M. J. (2010). What More Than Parental Income, Education and Occupation? An Exploration of What Swedish Siblings Get from Their Parents. *The B.E. Journal of Economic Analysis & Policy*, 10(1).

Black, S. E., Grönqvist, E., and Öckert, B. (2018). Born to Lead? The Effect of Birth Order on Noncognitive Abilities. *Review of Economics and Statistics*, 100(2):274–286.

Bonett, D. G. (2007). Transforming Odds Ratios into Correlations for Meta-Analytic Research. *American Psychologist*, 62(3):254–255.

Eriksson, K. H., Hjalmarsson, R., Lindquist, M. J., and Sandberg, A. (2016). The Importance of Family Background and Neighborhood Effects as Determinants of Crime. *Journal of Population Economics*, 29(2):219–262.

Lindahl, L. (2011). A Comparison of Family and Neighborhood Effects on Grades, Test Scores, Educational Attainment and Income-Evidence from Sweden. *Journal of Economic Inequality*, 9(2):207–226.

Mazumder, B. (2008). Sibling Similarities and Economic Inequality in the US. *Journal of Population Economics*, 21(3):685–701.

Mishkin, E. (2017). Gender and Sibling Dynamics in the Intergenerational Transmission of Entrepreneurship. *Working Paper*.

Raaum, O., Salvanes, K. G., and Sørensen, E. Ø. (2006). The Neighbourhood Is Not What It Used to Be. *Economic Journal*, 116(1):200–222.

Solon, G. (1999). Intergenerational Mobility in the Labor Market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3, pages 1761–1800. Elsevier, Amsterdam.

Solon, G., Page, M. E., and Duncan, G. J. (2000). Correlations between Neighboring Children in Their Subsequent Educational Attainment. *Review of Economics and Statistics*, 82(3):383–392.

Vladasel, T. (2018). Same, but Different? Birth Order, Family Size, and Sibling Sex Composition Effects in Entrepreneurship. *SOFI Working Paper*, 8/2018.