

Table S1. Study design, sample size, sample quality control, and self-employment measure within each study.

Study		Study design	Total sample size (<i>n</i>)	Sample QC		Sample size after QC (<i>n</i>)	Self-employment measure	References
Abbreviation	Full name			Call rate	Other exclusions			
AGES	Age, Gene/Environment Susceptibility–Reykjavik Study	Population-based	3,219	≥ 98%	1) Mismatch previous genotypes; 2) gender discrepancy; 3) missing self-employment status.	3,219	Questionnaire: Are you a employer or manager?	1) Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, et al. (2007) Age, Gene/Environment Susceptibility–Reykjavik Study: Multidisciplinary applied phenomics. <i>Am J Epidemiol</i> 165: 1076–1087.
ASPS	Austrian Stroke Prevention Study	Population-based	2,008	≥ 98%	1) Excess autosomal heterozygosity; 2) mismatch between called and phenotypic gender; 3) outliers in the IBD analysis; 4) missing self-employment status.	834	Participants were systematically asked about their life-long professional activities and this information is recorded in the subjects' charts. In addition information was gathered as to whether the subjects conducted their professional activities as employees or self-employed.	1) Schmidt R, Lechner H, Fazekas F, Niederkorn K, Reinhart B, et al. (1994) Assessment of cerebrovascular risk profiles in healthy persons: definition of research goals and the Austrian Stroke Prevention Study (ASPS). <i>Neuroepidemiology</i> 13: 308–313.
ERF	Erasmus Rucphen Family study	Family- based	3,500	≥ 98%	1) Heterozygosity: FDR<1%; 2) ethnic outliers; 3) duplicates; 4) gender mismatch; 5) excess IBS incompatible with pedigree; 6) missing self-employment status.	1,071	Individuals were asked if they were ever self-employed (no; yes—once; yes—more than once; always; not applicable/no response) and coded as self-employed if they were self-employed at at least than once. The control group consists of people who were never self-employed.	1) Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS (2005) The effect of genetic drift in a young genetically isolated population. <i>Ann Hum Genet</i> 69: 288–295.
GHS	Gutenberg Health Study	Population-based	3,500	≥ 97%	1) Deviations in expected heterozygosity; 2) doubtful IBS patterns; 3) missing self-employment status.	3,130	Categorization of the actual job or last job (for participants who retired); questions on each job phase of the professional career were collected in a computer-assisted personal interview (CAPI).	1) Rotival M, Zeller T, Wild PS, Maouche S, Szymczak S, et al. (2011) Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. <i>PLoS Genet</i> 7: e1002367.
H2000	Health 2000	Population-based	8,028	≥ 95%	1) Excess heterozygosity; 2) relatedness and/or failed gender check; 3) missing self-employment status.	2,123	Participants were asked “Are you/were you 1) Salary earner, 2) Agricultural entrepreneur, 3) Other entrepreneur, 4) Self-employed person or freelancer, 5) Working in a family member's farm without a salary, 6) Working in a family member's business without a salary, 7) Other, 8) Never been working in a full time job, 9) don't know” Salary earners were coded as controls and participants that answered 2,3 or 4 were coded as cases. Other participants were excluded.	1) Aromaa A, Koskinen,S. (Eds) (2004). Health and functional capacity in Finland. Publications of the National Public Health Institute, KTL B12: Helsinki: Finland.

HBCS	Helsinki Birth Cohort Study	Birth cohort study	8,760	≥ 95%	1) Excess heterozygosity; 2) relatedness and/or failed gender check; 3) missing self-employment status.	1,724	In the HBCS, data on classification of the socio-economic status and self-employment were derived from the Statistics Finland. These data were available with five-year-interval from 1970 to 2000. Participant were classified as self-employed if at any of the time points they indicated self-employment. Participants were excluded if data in more than two (out of seven) time points were missing ($n = 4$). In addition, those controls with indication of self-employment were excluded.	1) Eriksson JG (2005) Early growth and adult health outcomes—lessons learned from the Helsinki Birth Cohort Study. <i>Matern Child Nutr</i> 1: 149–154.
HRS	Health and Retirement Study	Population-based	12,507	≥ 98%	1) Relatedness check; 2) ethnic outliers; 3) missing self-employment status.	6,220	From the HRS RAND v.L dataset the binary variables $r*slfemp$ were used that indicate if an individual was self-employed or working for someone else in wave *. Individuals are coded as self-employed if they responded in at least one data wave to be self-employed. The control group consists of people who were never self-employed but indicated at least once to work for someone else.	1) Weir D (2007) <i>Elastic Powers: The Integration of Biomarkers into the Health and Retirement Study</i> , Washington, D.C, USA: The National Academies Press, chapter Biosocial Surveys. pp. 78–95.
KORA S4	Cooperative Health Research in the Region of Augsburg	Population-based	4,261	> 93%	1) Gender mismatch; 2) missing self-employment status.	1,724	Questionnaire based. “Which position do you/did you have in your job?”	1) Wichmann HE, Gieger C, Illig T, MONICA/KORA Study Group (2005) Kora-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. <i>Gesundheitswesen</i> 67 Suppl 1: S26–S30.
NFBC1966	Northern Finland Birth Cohort 1966	Population-based	12,231	≥ 95%	1) Gender discrepancy with genetic data from X-linked markers; 2) withdrawn consent; 3) duplicates and first and second degree relatives; 4) contaminated samples. 5) missing self-employment status.	4,234	Based on questionnaire questions Q1-Q3, the subjects were classified into 9 groups according to the instructions from Statistics Finland (Reference: Tilastokeskus Käsikirjoja 17: Sosioekonomisen aseman luokitus 1989): 1) Farm businessmen 2) Other entrepreneurs 3) Upper white-collar workers 4) Lower white-collar workers 5) Blue-collar workers 6) Students 7) Pensioners 8) Unemployed (incl. long-term unemployed and unclassified) 9) Socio-economic status unknown. For the present analysis groups 1) and 2) are considered as cases and others as controls, excluding group 7) pensioners and from group 8) long-term unemployed and unclassified.	1) Rantakallio P (1969) Groups at risk in low birth weight infants and perinatal mortality. <i>Acta Paediatr Scand</i> 193: Suppl 193:1–71 2) Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, et al. (2009) Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. <i>Nat Genet</i> 41: 35–46.

NTR1	Netherlands Twin Register Cohort 1	Twin study	29,852	> 90%	1) Presence of genetic data; 2) gender discrepancy with genetic data; 3) unexpected IBS sharing; 4) contaminated samples; 5) duplicates and first and second degree relatives; 6) missing self-employment status.	1,555	Data came from eight surveys. Participants were asked to indicate whether they were self-employed (1991, 1993, 1995, 2004, 2009) or to indicate the type of organisation that they worked in, with being self-employed as one of the answer categories (1997, 2000, 2002). If they indicated to be self-employed in any one of the surveys, they were classified as self-employed.	1) Boomsma DI, de Geus EJC, Vink JM, Stubbe JH, Distel MA, et al. (2006) Netherlands Twin Register: From twins to twin families. <i>Twin Res Hum Genet</i> 9: 849–857. 2) Boomsma DI, Willemsen G, Sullivan PF, Heutink P, Meijer P, et al. (2008) Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. <i>Eur J Hum Genet</i> 16: 335–342.
NTR2	Netherlands Twin Register Cohort 2	Twin study	29,852	> 90%	1) Presence of genetic data; 2) gender discrepancy with genetic data; 3) unexpected IBS sharing; 4) contaminated samples; 5) duplicates and first and second degree relatives; 6) missing self-employment status.	984	Data came from eight surveys. Participants were asked to indicate whether they were self-employed (1991, 1993, 1995, 2004, 2009) or to indicate the type of organisation that they worked in, with being self-employed as one of the answer categories (1997, 2000, 2002). If they indicated to be self-employed in any one of the surveys, they were classified as self-employed.	1) Willemsen G, de Geus EJC, Bartels M, van Beijsterveldt CEMT, Brooks AI, et al. (2010) The Netherlands Twin Register biobank: A resource for genetic epidemiological studies. <i>Twin Res Hum Genet</i> 13: 231–245.
RS-I	Rotterdam Study Baseline	Population-based	7,983	≥ 97.5%	1) Gender mismatch with typed Xlinked markers; 2) excess autosomal heterozygosity > 0.336–FDR>0.1%; 3) duplicates and/or 1st or 2nd degree relatives using IBS probabilities >97% from PLINK; 4) ethnic outliers using IBS distances > 3SD from PLINK; 5) missing self-employment status.	5,374	Detailed background information on the entire work-life history of all participants was available, including the number and duration of self-employment spells. Participants who had at least one spell in self-employment were coded as cases, those with zero spells as controls. Participants with incomplete work-life histories were excluded.	1) Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HLA, et al. (2011) The rotterdam study: 2012 objectives and design update. <i>Eur J Epidemiol</i> 26: 657–686. 2) Estrada K, Abuseiris A, Grosveld FG, Uitterlinden AG, Knoch TA, et al. (2009) Grimp: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. <i>Bioinformatics</i> 25: 2750–2752.
RS-II	Rotterdam Study Extension of Baseline	Population-based	3,011	≥ 97.5%	1) Gender mismatch with typed Xlinked markers; 2) excess autosomal heterozygosity (F<-0.055); 3) duplicates and/or 1st degree relatives using IBD PiHAT >40% from PLINK; 4) ethnic outliers IBS distances > 4SD mean HaMAP CEU cluster from PLINK; 5) missing self-employment status.	2,066	Information on current employment status was available and self-employed participants were coded as cases. Participants that indicated employment or that were collaborating family members were coded as controls. Other participants were excluded.	1) Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HLA, et al. (2011) The rotterdam study: 2012 objectives and design update. <i>Eur J Epidemiol</i> 26: 657–686. 2) Estrada K, Abuseiris A, Grosveld FG, Uitterlinden AG, Knoch TA, et al. (2009) Grimp: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. <i>Bioinformatics</i> 25: 2750–2752.

RS-III	Rotterdam Study Young	Population-based	3,932	≥ 97.5%	1) Gender mismatch with typed Xlinked markers; 2) excess autosomal heterozygosity ($F < -0.055$); 3) duplicates and/or 1st degree relatives using IBD PiHAT >40% from PLINK; 4) ethnic outliers IBS distances > 4SD mean HaMAP CEU cluster from PLINK; 5) missing self-employment status.	1,925	Information on current or last employment status (if retired) was available and self-employed participants were coded as cases. Participants that indicated employment or that were collaborating family members were coded as controls. Other participants were excluded.	1) Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HLA, et al. (2011) The rotterdam study: 2012 objectives and design update. <i>Eur J Epidemiol</i> 26: 657–686. 2) Estrada K, Abuseiris A, Grosveld FG, Uitterlinden AG, Knoch TA, et al. (2009) Grimp: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. <i>Bioinformatics</i> 25: 2750–2752.
SardinIA	SardiNIA Study of Aging	Population-based	6,148	≥ 95%	1) missing genotype; 2) missing self-employment status.	4,142	Information on current employment status was available and self-employed participants were coded as cases.	1) Pilia G, Chen WM, Scuteri A, Orr M, Albai G, et al. (2006) Heritability of cardiovascular and personality traits in 6,148 Sardinians. <i>PLoS Genet</i> 2: e132.
SHIP	Study of Health in Pomerania	Population-based	4,308	≥ 92%	1) Duplicate samples (by IBS); 2) reported/genotyped gender mismatch; 3) missing self-employment status.	4,063	“Which occupational position do you currently have?” Participants were coded as self-employed if they answered that they were farmers with more than 10 hectare property, university graduates with a liberal profession (physician, lawyer, tax accountant, etc.) or self-employed (in business, craft, or the tertiary sector).	1) Völzke H, Alte D, Schmidt CO, Radke D, Lohr R, et al. (2011) Cohort profile: The Study of Health in Pomerania. <i>Int J Epidemiol</i> 40: 294–307.
STR	Swedish Twin Registry	Twin study	10,946	≥ 97%	1) Sex-check (heterozygosity of X-chromosomes); 2) deviations in heterozygosity of more than 5 SD from the population mean; 3) cryptically relatedness check; 4) missing self-employment status.	3,271	Participants were asked “(1) have you ever run a business?” and “(2) how many businesses have you in total been part of starting?” Participants were coded as cases if they answered “Yes” to the first and “≥1” to the second question. Controls were those that answered “No” to the first and “0” to the second question. Other participants were excluded.	1) Lichtenstein P, Sullivan PF, Cnattingius S, Gatz M, Johansson S, et al. (2006) The Swedish Twin Registry in the third millennium: An update. <i>Twin Res Hum Genet</i> 9: 875–882. 2) Cesarini D, Johannesson M, Magnusson PKE, Wallace B (2011) The behavioral genetics of behavioral anomalies. <i>Manage Sci</i> : In press.
THISEAS	The Hellenic study of Interactions between SNPs & Eating in Atherosclerosis Susceptibility	CAD case-control	1,877	> 95%	1) Missing self-employment status; 2) heterozygosity; 3) gender mismatch; 4) ethnic outliers.	685	Information on current employment status was available. Participants were asked whether they were: 1) civil servant, 2) private employee, 3) self-employed, 4) part-time employee, 5) retired, or 6) house holding. Self-employed participants were coded as cases. Participants who were retired were excluded.	1) Theodoraki EV, Nikopentis T, Suhorutsenko J, Peppas V, Fili P, et al. (2010) Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. <i>BMC Med Genet</i> 11: 28.

TwinsUK	The UK Adult Twin Registry	Twin study	4,427	≥ 95%	<p>1) Heterozygosity across all SNPs ≥ 2 s.d. from the sample mean;</p> <p>2) evidence of non-European ancestry as assessed by PCA comparison with HapMap3 populations;</p> <p>3) observed pairwise IBD probabilities suggestive of sample identity errors;</p> <p>4) missing self-employment status.</p>	3,155	Information on the duration of self-employment was available. Participants that indicated to have spent some time in self-employment were coded as cases, those that indicated to have never been self-employed were coded as controls. Other participants were excluded.	<p>1) Moayyeri A, Hammond CJ, Valdes AM, Spector TD (2012) Cohort profile: TwinsUK and healthy ageing twin study. <i>Int J Epidemiol</i>: In press.</p> <p>2) Aulchenko YS, Ripke S, Isaacs A, van Duijn CM (2007) GenABEL: An R library for genome-wide association analysis. <i>Bioinformatics</i> 23: 1294–1296.</p>
YFS	The Cardiovascular Risk in Young Finns Study	Population-based cohort	3,596	≥ 95%	<p>1) Excess heterozygosity;</p> <p>2) relatedness and/or failed gender check.</p>	2,358	Study subjects were asked in years 1986, 1989, 1992 and 2001 if they were 1) Salary earners 2) Farm businessmen 3) Unpaid workers at their family farm 4) Entrepreneurs 5) Unpaid workers at their family business or 6) Others. Persons who indicated to be farm businessmen or entrepreneurs in any of the surveys were classified as cases. Those who hadn't in any time point reported to be an entrepreneur or farm businessman were selected as controls.	<p>1) Raitakari OT, Juonala M, Rönnemaa T, Keltikangas-Järvinen L, Räsänen L, et al. (2008). Cohort profile: The cardiovascular risk in Young Finns Study. <i>Int J Epidemiol</i> 37: 1220–1226.</p>

Table S2. Genotyping, imputation, SNP quality control, and statistical analysis within each study.

Genotyping			SNP inclusion criteria				Imputation and quality control before meta-analysis				Association analysis	
Study	Platform(s), SNP panel(s)	Genotyping calling algorithm	MAF	Call rate	HWE	Genotyped SNPs after QC	Imputation software	SNP inclusion criteria		Imputed SNPs after QC	Software	Covariates
								MAF	Imputation quality			
AGES	Illumina Human370CNV	BeadStudio	$\geq 1\%$	$\geq 98\%$	$\geq 10^{-6}$	317,344	MACH	$\geq 1\%$	$Rsq \geq 0.4$	2,176,303	ProbABEL	1) Sex in pooled sample; 2) first four PCs in all samples; 3) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50 .
ASPS	Illumina Human610-Quad BeadChip	BeadStudio	$\geq 1\%$	$\geq 98\%$	$\geq 10^{-6}$	550,635	MACH	$\geq 5\%$	$Rsq \geq 0.4$	2,164,654	GenABEL	1) Sex in pooled sample; 2) first four PCs.
ERF	Illumina 6K, 318K, 370K, Affymetrix 250K, Illumina 610K	Beadstudio using Hap300v2 cluster file	$\geq 1\%$	$\geq 98\%$	$\geq 10^{-6}$	650,197	MACH	$\geq 1\%$	$Rsq \geq 0.4$	2,353,164	ProbABEL	1) Sex in pooled sample; 2) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50 .
GHS	Affymetrix 6.0	Birdseed	$\geq 1\%$	$\geq 98\%$	$\geq 10^{-4}$	649,182	IMPUTE	$\geq 5\%$	$info \geq 0.4$	2,220,912	SNPTEST	1) Sex in pooled sample; 2) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50
Health 2000	Illumina Human610-Quad BeadChip	Illuminus	$\geq 1\%$	$\geq 95\%$	$\geq 10^{-6}$	555,418	MACH	$\geq 1\%$	$Rsq \geq 0.4$	2,463,699	ProbABEL	1) Sex in pooled sample; 2) first four PCs; 3) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50 .
HBCS	Illumina Human660-Quad BeadChip	Illuminus	$\geq 1\%$	$\geq 95\%$	$\geq 10^{-6}$	533,491	MACH	$\geq 1\%$	$Rsq \geq 0.4$	2,416,556	ProbABEL	1) Sex in pooled sample; 2) first four PCs in all samples.
HRS	Illumina Omni2.5 Beadchip	GenomeStudio	$\geq 1\%$	$\geq 98\%$	$\geq 10^{-4}$	2,195,306	MACH	$\geq 1\%$	$Rsq \geq 0.4$	2,227,690	PLINK	1) Sex in pooled sample; 2) first four PCs in all samples; 3) age dummies for the categories 30–39 (reference), 40–49, ≥ 50 .
KORA S4	Affymetrix 6.0	Birdseed2	—	—	—	909,622	IMPUTE	$\geq 1\%$	$info \geq 0.4$	2,521,850	SNPTEST	1) Sex in pooled sample; 2) age categories (≤ 39 ; 40–49; ≥ 50).
NFBC1966	Illumina HumanCNV-370DUO Analysis BeadChip	Beadstudio	$\geq 5\%$	$\geq 95\%$	$\geq 10^{-4}$	328,007	IMPUTE	$\geq 1\%$	$info \geq 0.4$	2,405,775	SNPTEST	1) Sex in pooled sample; 2) first four PCs.
NTR1	Perlegen–Affymetrix 600K	Affymetrix Proprietary	$\geq 1\%$	$\geq 95\%$	$> 10^{-5}$	427,049	IMPUTE	$\geq 1\%$	$info \geq 0.4$	2,420,149	SNPTEST	1) Sex in pooled sample; 2) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50 .
NTR2	IlluminaTM Human660W-QuadXX	Beadstudio	$> 1\%$	$> 95\%$	$> 10^{-5}$	528,072	IMPUTE	$\geq 1\%$	$info \geq 0.4$	2,532,400	SNPTEST	1) Sex in pooled sample; 2) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50 .
RS-I	Illumina HumanHap 550 V.3	BeadStudio Genecall	$\geq 1\%$	$\geq 98\%$	$\geq 10^{-6}$	512,349	MACH	$\geq 1\%$	$Rsq \geq 0.4$	2,433,150	MACH2DAT	1) Sex in pooled sample; 2) first four PCs.
RS-II	Illumina HumanHap 550 V.3 DUO; Illumina HumanHap 610 QUAD	Genomestudio Genecall	$\geq 1\%$	$\geq 97.5\%$	$\geq 10^{-6}$	466,389	MACH	$\geq 1\%$	$Rsq \geq 0.4$	2,432,613	MACH2DAT	1) Sex in pooled sample; 2) first four PCs.

RS-III	Illumina HumanHap 610 QUAD	Genomestudio Genecall	≥ 1%	≥ 97.5%	≥ 10 ⁻⁶	514,073	MACH	≥ 1%	Rsq ≥ 0.4	2,436,797	MACH2DAT	1) Sex in pooled sample; 2) first four PCs; 3) dummy for age ≥ 50.
SardinIA	Affymetrix 10k, 500k, 1M	BRLMM	≥ 5%	≥ 95%	≥ 10 ⁻⁶	356,359	MACH	≥ 5%	Rsq ≥ 0.4	1,972,533	Merlin	1) Sex in pooled sample.
SHIP	Affymetrix Human SNP Array 6.0	Birdseed V2	—	—	—	869,224	IMPUTE	≥ 1%	info ≥ 0.4	2,514,047	QUICKTEST	1) Sex in pooled sample; 2) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50.
STR	Illumina HumanOmniExpress BeadChip	GenomeStudio Genecall	≥ 1%	≥ 97%	≥ 10 ⁻⁷	644,556	IMPUTE	≥ 1%	info ≥ 0.4	2,481,423	Merlin	1) Sex in pooled sample; 2) first four PCs.
THISEAS	CardioMetaboChip Array	GenoSNP	≥ 5%	≥ 95%	≥ 10 ⁻⁶	96,015	—	≥ 1%	—	95,510 ^a	PLINK	1) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50.
TwinsUK	Illumina Human610-Quad BeadChip	Illuminus	≥ 5%	≥ 95%	≥ 10 ⁻⁵	536,559	IMPUTE	≥ 1%	info ≥ 0.4	2,262,054	GenABEL	1) Sex in pooled sample; 2) first four PCs; 3) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50.
YFS	Illumina Human670-Quad Custom BeadChip	Illuminus	≥ 1%	≥ 95%	≥ 10 ⁻⁶	546,677	MACH	≥ 1%	Rsq ≥ 0.4	2,409,746	ProbABEL	1) Sex in pooled sample; 2) first four PCs in all samples; 3) age dummies for the categories ≤ 29 (reference), 30–39, 40–49, ≥ 50.

^a Number of genotyped SNPs after filtering on minor allele frequency ≥ 1%.

Table S3. Genomic inflation factors.

Study	Pooled	Males	Females
AGES	1.000	1.019	0.992
ASPS	1.117	0.998	0.998
ERF	1.105	1.099	1.042
GHS	0.996	1.006	0.999
H2000	0.990	0.994	0.995
HBCS	1.005	0.995	1.007
HRS	1.008	1.006	0.998
KORA S4	1.000	1.008	1.001
NFBC1966	1.003	1.007	1.004
NTR1	1.006	1.015	1.002
NTR2	1.004	1.021	1.022
RS-I	1.022	1.008	1.003
RS-II	1.008	1.015	0.998
RS-III	1.000	1.004	1.013
SardinIA	1.074	1.038	1.156
SHIP	1.010	1.021	0.989
STR	0.996	0.997	0.990
THISEAS ^a	—	—	—
TwinsUK ^b	1.015	—	1.011
YFS	1.010	1.015	1.007

^a THISEAS did not provide results using genome-wide SNP data and genomic control lambdas were therefore not estimated.

^b The number of male subjects was insufficient for a male stratified analysis.

Table S4. Replication results of the twenty suggestive SNPs ($p < 1 \times 10^{-5}$) from the self-employment discovery meta-analyses for pooled males and females.

Discovery meta-analysis											Swedish Twin Registry		Combined meta-analysis		Improvement?
SNP	Chr.	Pos.	Effect / non-effect allele	Nearest gene	Distance to gene (bp)	I^2	Cochran's Q test p -value	n	p -value	Overall freq.	p -value	Freq.	p -value	Direction	
rs6906622	6	18,596,287	T/C	<i>RNF144B</i>	19,460	0.0	0.644	49,942	4.10×10^{-6}	0.21	0.30	0.19	2.54×10^{-5}	+++++?+++	no
rs477588	6	18,586,095	T/C	<i>RNF144B</i>	9,268	0.0	0.646	49,942	4.15×10^{-6}	0.21	0.32	0.19	2.46×10^{-5}	+++++?+++	no
rs9350075	6	18,595,091	A/C	<i>RNF144B</i>	18,264	0.0	0.640	49,942	4.26×10^{-6}	0.79	0.31	0.81	2.63×10^{-5}	-----?---	no
rs3920593	6	18,595,112	T/C	<i>RNF144B</i>	18,285	0.0	0.641	49,942	4.35×10^{-6}	0.79	0.30	0.81	2.68×10^{-5}	-----?---	no
rs6928004	6	18,597,621	A/G	<i>RNF144B</i>	20,794	0.0	0.640	49,942	4.47×10^{-6}	0.79	0.31	0.81	2.71×10^{-5}	-----?---	no
rs4716268	6	18,598,474	A/C	<i>RNF144B</i>	21,647	0.0	0.639	49,942	4.60×10^{-6}	0.21	0.31	0.19	2.76×10^{-5}	+++++?+++	no
rs10776614	10	49,433,172	T/C	<i>ARHGAP22</i>	49,972	0.0	0.980	49,942	4.79×10^{-6}	0.16	0.78	0.17	6.78×10^{-6}	-----?---	no
rs2358531	5	75,515,542	A/G	<i>SV2C</i>	100,482	0.0	0.660	46,812	4.79×10^{-6}	0.71	0.01	0.72	1.42×10^{-4}	---?-----?---	no
rs747533	10	49,426,264	A/G	<i>ARHGAP22</i>	56,880	0.0	0.973	49,942	5.26×10^{-6}	0.26	0.92	0.26	1.14×10^{-5}	-----?---	no
rs510367	6	18,591,491	A/G	<i>RNF144B</i>	14,664	0.0	0.677	49,942	5.33×10^{-6}	0.21	0.31	0.19	3.20×10^{-5}	+++++?+++	no
rs6899634	6	18,600,965	A/C	<i>RNF144B</i>	24,138	0.0	0.619	49,942	5.54×10^{-6}	0.21	0.32	0.19	3.19×10^{-5}	+++++?+++	no
rs17166082	7	131,363,900	A/G	<i>PLXNA4</i>	94,730	0.0	0.821	39,758	5.82×10^{-6}	0.06	0.88	0.05	9.37×10^{-6}	-?-?---?---	no
rs994208	14	33,531,622	C/G	<i>EGLN3</i>	41,587	0.0	0.741	49,942	6.11×10^{-6}	0.66	0.62	0.68	2.05×10^{-5}	-----?---	no
rs3847697	12	57,282,257	T/C	<i>LRIG3</i>	269,947	0.0	0.762	41,566	6.79×10^{-6}	0.44	0.64	0.43	2.19×10^{-5}	-----?---?---	no
rs9371065	6	18,601,109	A/T	<i>RNF144B</i>	24,282	0.0	0.682	49,942	8.84×10^{-6}	0.22	0.32	0.19	4.84×10^{-5}	+++++?+++	no
rs2057556	6	18,602,356	T/C	<i>RNF144B</i>	25,529	0.0	0.693	49,942	8.98×10^{-6}	0.78	0.33	0.81	4.88×10^{-5}	-----?---	no
rs10514062	5	75,513,972	A/T	<i>SV2C</i>	98,912	0.0	0.636	49,942	8.99×10^{-6}	0.72	0.01	0.72	2.44×10^{-4}	-----?---	no
rs3742467	14	49,709,284	T/C	<i>SOS2</i>	55,689	3.5	0.413	43,722	9.11×10^{-6}	0.88	0.44	0.89	3.97×10^{-5}	++++-?++++?+++	no
rs1324564	6	18,603,428	T/C	<i>RNF144B</i>	26,601	0.0	0.698	49,942	9.21×10^{-6}	0.78	0.33	0.81	4.97×10^{-5}	-----?---	no
rs1359339	6	18,606,507	T/G	<i>RNF144B</i>	29,680	0.0	0.775	49,942	9.77×10^{-6}	0.22	0.33	0.19	5.24×10^{-5}	+++++?+++	no

Chr.: chromosome; Pos.: position; Overall freq.: average effect allele frequency; In the column "direction", the studies are in the following order: 1. AGES, 2. ASPS, 3. ERF, 4. GHS, 5. H2000, 6. HBCS, 7. HRS, 8. KORA S4, 9. NFBC1966, 10. NTR1, 11. NTR2, 12. RS-I, 13. RS-II, 14. RS-III, 15. SardinIA, 16. SHIP, 17. THISEAS, 18. TwinsUK, 19. YFS, 20. STR; A question mark indicates that the SNP was not tested in that specific study.

Table S5. Replication results of the 22 suggestive SNPs ($p < 1 \times 10^{-5}$) from the self-employment discovery meta-analyses for males only.

Discovery meta-analysis											Swedish Twin Registry		Combined meta-analysis		
SNP	Chr.	Pos.	Effect / non-effect allele	Nearest gene	Distance to gene (bp)	I^2	Cochran's Q test p -value	n	p -value	Overall freq.	p -value	Freq.	p -value	Direction	Improvement?
rs6738407	2	196,851,876	A/G	HECW2	79,655	0.0	0.819	21,508	1.52×10^{-7}	0.20	0.26	0.20	7.93×10^{-8}	-----+-----?--	yes
rs7558876	2	196,884,193	T/C	HECW2	111,972	0.0	0.676	21,508	2.02×10^{-7}	0.25	0.98	0.26	4.85×10^{-7}	-----+-----?--	no
rs4850390	2	196,851,234	A/T	HECW2	79,013	0.0	0.807	21,508	2.18×10^{-7}	0.20	0.26	0.20	1.14×10^{-7}	-----+-----?--	yes
rs12619045	2	196,883,649	A/G	HECW2	111,428	0.0	0.707	21,508	2.41×10^{-7}	0.75	0.97	0.74	5.81×10^{-7}	+++++++-----+?+-	no
rs17829296	2	196,871,827	A/G	HECW2	99,606	0.0	0.643	21,508	2.75×10^{-7}	0.80	0.77	0.81	9.04×10^{-7}	+++++++-----+?+-	no
rs1869794	2	196,880,651	T/G	HECW2	108,430	0.0	0.650	21,508	2.82×10^{-7}	0.20	0.81	0.19	8.72×10^{-7}	-----+-----?--	no
rs4145361	2	196,849,516	T/C	HECW2	77,295	0.0	0.767	21,508	3.94×10^{-7}	0.80	0.27	0.80	2.12×10^{-7}	+++++++-----+?++	yes
rs11889112	2	196,873,346	A/G	HECW2	101,125	0.0	0.624	21,508	4.23×10^{-7}	0.20	0.81	0.19	1.27×10^{-6}	-----+-----?--	no
rs11901261	2	196,874,778	A/C	HECW2	102,557	0.0	0.623	21,508	4.28×10^{-7}	0.80	0.81	0.81	1.28×10^{-6}	+++++++-----+?+-	no
rs11902488	2	196,875,256	A/G	HECW2	103,035	0.0	0.623	21,508	4.31×10^{-7}	0.80	0.81	0.81	1.29×10^{-6}	+++++++-----+?+-	no
rs10190368	2	196,870,691	A/G	HECW2	98,470	0.0	0.618	21,508	4.50×10^{-7}	0.80	0.87	0.81	1.22×10^{-6}	+++++++-----+?+-	no
rs6757097	2	196,882,699	T/C	HECW2	110,478	0.0	0.705	21,508	4.87×10^{-7}	0.25	0.95	0.26	1.16×10^{-6}	-----+-----?--	no
rs2889155	2	196,852,550	T/G	HECW2	80,329	0.0	0.704	18,680	5.27×10^{-7}	0.82	0.61	0.82	6.10×10^{-7}	+++++?-----+?++	no
rs2124384	2	196,877,435	T/C	HECW2	105,214	0.0	0.646	19,786	5.49×10^{-7}	0.80	0.83	0.81	1.58×10^{-6}	+++++++-----+?+-	no
rs12473229	2	196,878,519	T/G	HECW2	106,298	0.0	0.681	18,680	9.25×10^{-7}	0.75	0.94	0.74	2.16×10^{-6}	+++++?+-----+?+-	no
rs2305567	2	196,881,061	A/G	HECW2	108,840	0.0	0.690	21,508	9.84×10^{-7}	0.75	0.94	0.74	2.30×10^{-6}	+++++++-----+?+-	no
rs6730618	2	196,865,163	T/C	HECW2	92,942	0.0	0.720	21,508	3.34×10^{-6}	0.78	0.92	0.79	5.89×10^{-6}	+-----+-----+?++	no
rs6825440	4	183,636,063	A/T	ODZ3	153,933	24.1	0.175	21,508	4.25×10^{-6}	0.24	0.07	0.21	9.68×10^{-7}	+-----+-----?--	yes
rs7904494	10	72,056,694	A/T	PRF1	24,157	15.2	0.280	18,680	6.74×10^{-6}	0.78	0.32	0.80	3.79×10^{-5}	+-----?-----?--	no
rs1378626	2	196,864,466	T/C	HECW2	92,245	0.0	0.667	21,508	7.62×10^{-6}	0.21	0.54	0.20	7.10×10^{-6}	+-----+-----?--	yes
rs4867424	5	32,331,331	T/C	MTMR12	17,540	0.0	0.919	21,508	8.39×10^{-6}	0.49	0.83	0.50	2.00×10^{-5}	+-----+-----?--	no
rs2712008	4	38,752,396	T/G	KLHL5	11,942	0.0	0.815	18,680	9.94×10^{-6}	0.14	0.44	0.11	4.26×10^{-5}	+-----+-----+?+-	no

Chr.: chromosome; Pos.: position; Overall freq.: average effect allele frequency; In the column "direction", the studies are in the following order: 1. AGES, 2. ASPS, 3. ERF, 4. GHS, 5. H2000, 6. HBCS, 7. HRS,

8. KORA S4, 9. NFBC1966, 10. NTR1, 11. NTR2, 12. RS-I, 13. RS-II, 14. RS-III, 15. SardINIA, 16. SHIP, 17. THISEAS, 18. YFS, 19. STR; A question mark indicates that the SNP was not tested in that specific study.

Table S6. Replication results of the sixteen suggestive SNPs ($p < 1 \times 10^{-5}$) from the self-employment discovery meta-analyses for females only.

Discovery meta-analysis											Swedish Twin Registry		Combined meta-analysis		
SNP	Chr.	Pos.	Effect / non-effect allele	Nearest gene	Distance to gene (bp)	I^2	Cochran's Q test p -value	n	p -value	Overall freq.	p -value	Freq.	p -value	Direction	Improvement?
rs2331548	4	170,199,179	A/G	<i>CBR4</i>	31,182	0.0	0.913	24,319	1.93×10^{-6}	0.96	0.95	0.96	4.30×10^{-6}	??+?+++++?+++	no
rs521326	6	52,927,336	A/G	<i>GSTA4</i>	23,373	0.0	0.558	28,174	2.92×10^{-6}	0.61	0.48	0.63	1.33×10^{-5}	-----+---?--+	no
rs1022335	2	145,813,253	A/T	<i>ZEB2</i>	818,867	0.0	0.762	24,782	3.02×10^{-6}	0.37	0.48	0.36	1.37×10^{-5}	-----?-----+---?--+	no
rs487642	6	52,926,012	A/G	<i>GSTA4</i>	24,697	0.0	0.557	28,174	3.05×10^{-6}	0.39	0.47	0.38	1.42×10^{-5}	+++++-----+---?+++	no
rs575861	6	52,931,798	A/C	<i>GSTA4</i>	18,911	0.0	0.599	28,174	3.83×10^{-6}	0.61	0.50	0.63	1.63×10^{-5}	-----+---?--+	no
rs10753804	1	168,583,032	T/C	<i>SCYL1BP1</i>	184,861	0.0	0.848	24,782	3.92×10^{-6}	0.49	0.98	0.42	7.97×10^{-6}	-----?-----+---?--+	no
rs521664	6	52,930,496	T/C	<i>GSTA4</i>	20,213	0.0	0.593	28,174	3.95×10^{-6}	0.61	0.53	0.63	1.60×10^{-5}	-----+---?--+	no
rs594614	6	52,927,410	A/C	<i>GSTA4</i>	23,299	0.0	0.528	28,440	4.15×10^{-6}	0.39	0.48	0.38	1.83×10^{-5}	+++++-----+---+---	no
rs562487	5	78,442,190	A/G	<i>BHMT</i>	1,169	0.0	0.630	28,174	4.49×10^{-6}	0.48	0.50	0.50	4.01×10^{-6}	+++++-----+---+---+---	yes
rs9557259	13	99,031,403	T/C	<i>TM9SF2</i>	18,125	0.0	0.592	12,506	5.16×10^{-6}	0.06	0.34	0.03	2.92×10^{-5}	??-?+++?+++++?+++?++	no
rs1383043	4	123,562,066	A/G	<i>ADAD1</i>	8,323	24.9	0.167	25,279	6.05×10^{-6}	0.38	0.71	0.35	1.78×10^{-5}	---+-----+-----?+++	no
rs9578700	13	23,775,308	A/G	<i>SPATA13</i>	740	38.9	0.047	28,174	6.53×10^{-6}	0.67	0.34	0.70	3.65×10^{-5}	-+++-----+-----?+++	no
rs4481907	1	168,599,170	T/C	<i>SCYL1BP1</i>	168,723	0.0	0.835	28,174	7.88×10^{-6}	0.51	0.99	0.58	1.48×10^{-5}	+++++-----+---+---+---	no
rs947230	13	23,772,926	A/C	<i>SPATA13</i>	3,122	40.1	0.041	28,174	8.02×10^{-6}	0.67	0.33	0.70	4.39×10^{-5}	-+++-----+-----?+++	no
rs7684834	4	123,260,318	A/G	<i>KIAA1109</i>	50,889	31.7	0.097	28,174	8.69×10^{-6}	0.62	0.82	0.65	1.27×10^{-5}	++-++++-+++++---?--+	no
rs557302	5	78,413,090	A/G	<i>BHMT2</i>	7,939	0.0	0.716	28,174	8.77×10^{-6}	0.49	0.69	0.49	1.06×10^{-5}	-----+---+-----?---	no

Chr.: chromosome; Pos.: position; Overall freq.: average effect allele frequency; In the column "direction", the studies are in the following order: 1. AGES, 2. ASPS, 3. ERF, 4. GHS, 5. H2000, 6. HBCS, 7. HRS.

8. KORA S4, 9. NFBC1966, 10. NTR1, 11. NTR2, 12. RS-I, 13. RS-II, 14. RS-III, 15. Sardinia, 16. SHIP, 17. THISEAS, 18. TwinsUK, 19. YFS, 20. STR; A question mark indicates that the SNP was not tested in that specific study

Table S7. Gene-based p -values for the candidate entrepreneurship genes for pooled males and females, males only, and females only.

Gene	Pooled	Males	Females
<i>ADORA2A</i>	0.228	0.464	0.293
<i>ADRA2A</i>	0.007	0.011	0.183
<i>COMT</i>	0.528	0.999	0.073
<i>DDC</i>	0.334	0.758	0.604
<i>DRD1</i>	0.666	0.331	0.366
<i>DRD2</i>	0.749	0.843	0.786
<i>DRD3</i>	0.012	0.010	0.603
<i>DRD4</i>	0.483	0.366	0.221
<i>DRD5</i>	0.689	0.803	0.417
<i>DYX1C1</i>	0.384	0.164	0.347
<i>HTR1B</i>	0.892	0.975	0.511
<i>HTR1E</i>	0.953	0.518	0.597
<i>HTR2A</i>	0.079	0.030	0.685
<i>KIAA0319 (DYX2)</i>	0.324	0.477	0.419
<i>ROBO1</i>	0.554	0.692	0.435
<i>SLC6A3 (DAT1)</i>	0.679	0.645	0.627
<i>SNAP25</i>	0.118	0.209	0.888

Table S8. Gene-based *p*-values for the top 25 genes associated with self-employment in the discovery meta-analysis for pooled males and females.

Chr.	Gene	Number of SNPs	Start position	Stop position	<i>p</i>-value
11	<i>SLC15A3</i>	78	60,461,135	60,475,833	1.63×10^{-4}
11	<i>TMEM132A</i>	67	60,448,488	60,461,207	2.46×10^{-4}
11	<i>PRPF19</i>	78	60,414,777	60,430,632	2.61×10^{-4}
11	<i>TMEM109</i>	68	60,438,252	60,447,491	2.73×10^{-4}
11	<i>CD6</i>	109	60,495,690	60,544,424	3.05×10^{-4}
9	<i>FANCG</i>	65	35,063,834	35,070,013	3.15×10^{-4}
6	<i>FBXL4</i>	132	99,428,321	99,502,570	3.31×10^{-4}
9	<i>PIGO</i>	68	35,078,687	35,086,579	3.51×10^{-4}
7	<i>SLC26A5</i>	112	102,780,412	102,873,834	3.53×10^{-4}
11	<i>ZP1</i>	73	60,391,590	60,399,740	3.84×10^{-4}
9	<i>DNAJB5</i>	62	34,979,784	34,988,428	4.02×10^{-4}
9	<i>VCP</i>	68	35,046,064	35,062,739	4.03×10^{-4}
9	<i>C9orf131</i>	59	35,031,101	35,035,988	5.02×10^{-4}
9	<i>KIAA1539</i>	68	35,094,117	35,105,893	5.05×10^{-4}
9	<i>STOML2</i>	67	35,089,888	35,093,154	5.55×10^{-4}
14	<i>C14orf138</i>	90	49,645,099	49,653,047	5.63×10^{-4}
9	<i>KIAA1045</i>	76	34,948,191	34,972,541	6.03×10^{-4}
10	<i>SHOC2</i>	95	112,713,902	112,763,413	6.81×10^{-4}
5	<i>PRLR</i>	293	35,099,984	35,266,334	7.03×10^{-4}
6	<i>IHPK3</i>	205	33,797,420	33,822,660	7.24×10^{-4}
14	<i>LOC196913</i>	102	49,620,118	49,629,111	1.00×10^{-3}
14	<i>SOS2</i>	159	49,653,595	49,767,849	1.02×10^{-3}
6	<i>C6orf125</i>	181	33,773,323	33,787,482	1.05×10^{-3}
19	<i>TMEM190</i>	41	60,580,015	60,581,424	1.22×10^{-3}
15	<i>TMOD2</i>	138	49,831,101	49,889,635	1.26×10^{-3}

Table S9. Gene-based *p*-values for the top 25 genes associated with self-employment in the discovery meta-analysis for males only.

Chr.	Gene	Number of SNPs	Start position	Stop position	<i>p</i>-value
4	<i>TMEM156</i>	171	38,644,835	38,710,436	1.61×10^{-4}
4	<i>KLHL5</i>	146	38,723,053	38,800,224	3.21×10^{-4}
11	<i>SLCO2B1</i>	139	74,539,810	74,594,947	3.57×10^{-4}
5	<i>STARD4</i>	96	110,861,920	110,876,056	4.19×10^{-4}
8	<i>TMEM67</i>	50	94,836,268	94,899,523	4.48×10^{-4}
17	<i>TNFSF12-TNFSF13</i>	63	7,393,139	7,405,649	4.67×10^{-4}
1	<i>OR2M2</i>	89	246,409,910	246,410,954	4.71×10^{-4}
4	<i>ING2</i>	74	184,663,213	184,669,243	4.73×10^{-4}
17	<i>SAT2</i>	47	7,470,280	7,471,889	4.98×10^{-4}
17	<i>TNFSF12</i>	61	7,393,098	7,401,931	5.16×10^{-4}
17	<i>TNFSF13</i>	49	7,402,339	7,405,642	5.16×10^{-4}
17	<i>EIF4A1</i>	44	7,416,780	7,423,048	5.18×10^{-4}
17	<i>SENP3</i>	46	7,406,042	7,416,011	5.29×10^{-4}
1	<i>OR2M5</i>	75	246,375,072	246,376,011	6.39×10^{-4}
2	<i>HECW2</i>	496	196,772,221	197,165,580	7.41×10^{-4}
17	<i>CD68</i>	44	7,423,528	7,426,153	7.41×10^{-4}
17	<i>SHBG</i>	48	7,474,215	7,477,395	8.32×10^{-4}
1	<i>OR2M3</i>	98	246,432,992	246,433,931	8.35×10^{-4}
17	<i>SOX15</i>	46	7,432,221	7,434,212	9.16×10^{-4}
17	<i>MPDUI</i>	48	7,427,853	7,432,247	1.00×10^{-3}
2	<i>PAX3</i>	205	222,772,850	222,871,944	1.02×10^{-3}
8	<i>RBM12B</i>	35	94,812,903	94,822,400	1.07×10^{-3}
15	<i>TMOD2</i>	139	49,831,101	49,889,635	1.10×10^{-3}
17	<i>FXR2</i>	58	7,435,271	7,458,796	1.22×10^{-3}
19	<i>TMEM190</i>	42	60,580,015	60,581,424	1.28×10^{-3}

Table S10. Gene-based *p*-values for the top 25 genes associated with self-employment in the discovery meta-analysis for females only.

Chr.	Gene	Number of SNPs	Start position	Stop position	<i>p</i>-value
21	<i>PCP4</i>	227	40,161,216	40,223,192	4.70×10^{-5}
9	<i>MELK</i>	94	36,562,904	36,667,679	2.02×10^{-4}
2	<i>FLJ20160</i>	155	190,981,325	191,075,286	2.48×10^{-4}
5	<i>BHMT2</i>	107	78,401,338	78,421,031	2.66×10^{-4}
4	<i>ADADI</i>	62	123,519,617	123,570,389	2.98×10^{-4}
4	<i>KIAA1109</i>	93	123,311,207	123,503,357	3.76×10^{-4}
4	<i>IL2</i>	47	123,592,075	123,597,100	3.81×10^{-4}
5	<i>BHMT</i>	74	78,443,359	78,463,869	4.26×10^{-4}
15	<i>CSPG4</i>	54	73,753,717	73,792,244	5.15×10^{-4}
4	<i>IL21</i>	106	123,753,232	123,761,661	5.46×10^{-4}
5	<i>ACTBL2</i>	128	56,811,599	56,814,393	6.98×10^{-4}
5	<i>ACTBL2</i>	128	56,811,599	56,814,393	7.33×10^{-4}
8	<i>PCMTD1</i>	194	52,892,692	52,974,288	7.53×10^{-4}
15	<i>SNX33</i>	36	73,728,402	73,738,023	8.43×10^{-4}
15	<i>ODF3L1</i>	37	73,803,373	73,807,082	9.80×10^{-4}
15	<i>ODF3L1</i>	37	73,803,373	73,807,082	9.90×10^{-4}
2	<i>HIBCH</i>	203	190,777,604	190,892,804	1.09×10^{-3}
2	<i>GKN2</i>	121	69,025,867	69,033,606	1.15×10^{-3}
22	<i>C22orf30</i>	74	30,402,241	30,438,731	1.16×10^{-3}
5	<i>C5orf35</i>	92	56,240,859	56,248,770	1.17×10^{-3}
6	<i>STX11</i>	120	144,513,346	144,554,769	1.17×10^{-3}
6	<i>GPX6</i>	59	28,579,051	28,591,549	1.32×10^{-3}
2	<i>INPP1</i>	110	190,916,440	190,944,636	1.34×10^{-3}
7	<i>SKAP2</i>	333	26,673,212	26,870,866	1.44×10^{-3}
2	<i>MGC13057</i>	140	190,710,730	190,776,455	1.58×10^{-3}

Table S11. Meta-analysis association results for SNP rs1486011 for pooled males and females, males only, and females only.

Sample	Effect / non- effect allele	EAF	n	p-value	Direction
Pooled	C/G	0.074	45,800	0.011	+++++++-----+?+?+
Males	C/G	0.077	19,786	0.046	+-----+?+?+
Females	C/G	0.072	25,754	0.112	+-----+?+?+

EAF: average allele frequency; In the column "direction", the studies are in the following order: 1. AGES, 2. ASPS, 3. ERF, 4. GHS, 5. H2000, 6. HBCS, 7. HRS, 8. KORA S4, 9. NFBC1966, 10. NTR1, 11. NTR2, 12. RS-I, 13. RS-II, 14. RS-III, 15. SARDINIA, 16. SHIP, 17. THISEAS, 18. TWINSUK (pooled and female sample) / YFS (male sample), 19. YFS (pooled and female sample); A question mark indicates that the SNP was not tested in that specific study.

Table S12. Results of the prediction analyses in STR for pooled males and females, males only, and females only.

Threshold	Pooled				Males				Females			
	Coeff.	<i>p</i> -value	<i>R</i> ² (%)	AUC	Coeff.	<i>p</i> -value	<i>R</i> ² (%)	AUC	Coeff.	<i>p</i> -value	<i>R</i> ² (%)	AUC
$p_T < 0.01$	-0.293	0.898	0.095%	0.518	-0.088	0.667	0.023%	0.515	-0.424	0.924	0.252%	0.530
$p_T < 0.05$	-0.005	0.531	0.000%	0.490	0.055	0.148	0.136%	0.507	-0.037	0.692	0.031%	0.508
$p_T < 0.1$	0.076	0.210	0.038%	0.517	0.086	0.144	0.141%	0.511	0.023	0.418	0.005%	0.498
$p_T < 0.2$	0.183	0.107	0.091%	0.519	0.123	0.170	0.114%	0.513	-0.006	0.513	0.000%	0.500
$p_T < 0.3$	0.277	0.078	0.119%	0.521	0.127	0.226	0.070%	0.512	-0.027	0.545	0.002%	0.508
$p_T < 0.4$	0.382	0.057	0.148%	0.524	0.221	0.149	0.135%	0.519	0.062	0.417	0.005%	0.506
$p_T < 0.5$	0.460	0.057	0.148%	0.524	0.285	0.132	0.155%	0.520	0.100	0.388	0.010%	0.519
$p_T < 0.6$	0.591	0.041	0.178%	0.526	0.266	0.187	0.099%	0.517	0.000	0.500	0.000%	0.500
$p_T < 0.7$	0.067	0.045	0.170%	0.523	0.286	0.203	0.087%	0.516	0.018	0.485	0.000%	0.497
$p_T < 0.8$	0.077	0.041	0.179%	0.525	0.360	0.178	0.106%	0.518	0.007	0.446	0.002%	0.513
$p_T < 0.9$	0.088	0.039	0.184%	0.525	0.041	0.173	0.111%	0.520	0.010	0.433	0.003%	0.512
$p_T \leq 1.0$	0.098	0.039	0.184%	0.525	0.047	0.169	0.115%	0.520	0.011	0.435	0.003%	0.516

Prediction results are based on a logistic regression of self-employment on the score controlling for the number of non-missing genotypes. Approximately 120,000 directly genotyped SNPs were used to calculate the score in STR for 12 overlapping significance thresholds. The R^2 is the Nagelkerke pseudo- R^2 from the logistic regression. The p -value indicates the significance of the score coefficient (one-sided Wald test, since the alternative hypothesis is that the score is positively correlated with self-employment). The AUC is the area under the receiver operating characteristic curve.