DATA APPENDIX

Information on all demographic variables (age, gender, children, parents, employment and occupations) are drawn from a set of registries previously known as the Integrated Database for Labor Market Research (IDA). These registries combine high-accuracy information across more than 150 government registries.

Data on psychiatric patients are drawn from the LPSYDIAG registry. Data on prescriptions come from the LMDB registry.

Information on families, households and demographics are from the BEF, FAIN, and FAM, registries. Data on employment, unemployment, and earnings are drawn from the IDAN, IND, and IDAP registries.

We link individual-level variables across these datasets using social security numbers (SSN). People born in Demark receive their SSNs at birth. Immigrants and foreign employees are assigned an SSN by the municipal office or the International Citizen Service when they receive a work permit or residence.

TABLE A1: LIST OF VARIABLES

Variable	Variable name	Definition	Years available	Registry name	Registry
Prescriptions and Diagnoses					
BD		Indicator for individuals with diagnosis code ICD-10: F31 and ICD-10: F30	1995-2015	Landspatientregistret for Psykiatri Diagnoser	LPSYDIAG
Schizophrenia		Indicator for individuals with diagnosis codes ICD-10: F20-F29	1995-2015	Landspatientregistret for Psykiatri Diagnoser	LPSYDIAG
Depression		Indicator for individuals with diagnosis codes ICD-10: F32	1995-2015	Landspatientregistret for Psykiatri Diagnoser	LPSYDIAG
Lithium		Indicator for individuals with at least 1 prescription of lithium (ATC: N05AN)	1995-2015	Medicinal Product Statistics	LMDB
Labor Market Variables					
Earnings	ERHVERVSINDK + NETOVSKUD	Sum of total wages for all jobs and income from self-employment	1995-2015	Income and Employment	IND, IDAP and IDAN
Disability	PSTILL	Indicator for individuals with variable PSTILL = 93	1995-2015	Demographics	IDAP
Days of unemployment	ARLEDGR	Number of days of unemployment (based on information from the unemployment funds)	1995-2015	Demographics	IDAP
Family					
Mother ID		Individual identifier of mother	1995-2015	Family information	BEF, FAIN and FAM

TABLE A2: DESCRIPTION OF DIAGNOSES

Variable	ICD code	ICD definitions
BD	ICD-10 30	A disorder characterized by two or more episodes in which the patient's mood and activity levels are significantly disturbed, this disturbance consisting in some occasions of an elevation of mood and increased energy and activity (hypomania or mania) and on others of a lowering of mood and decreased energy and activity (depression). Repeated episodes of hypomania or mania only are classified as bipolar.
Mania	ICD-10 31	A disorder which is elevated out of keeping with the patient's circumstances and may vary from carefree joviality to almost uncontrollable excitement. Elation is accompanied by increased energy, resulting in overactivity, pressure of speech, and a decreased need for sleep. Attention cannot be sustained, and there is often marked distractibility. Self-esteem is often inflated with grandiose ideas and overconfidence. Loss of normal social inhibitions may result in behavior that is reckless, foolhardy, or inappropriate to the circumstances, and out of character.
Depression	ICD-10: F32	A mental condition marked by ongoing feelings of sadness, despair, loss of energy, and difficulty dealing with normal daily life. Other symptoms of depression include feelings of worthlessness and hopelessness, loss of pleasure in activities, changes in eating or sleeping habits, and thoughts of death or suicide.
Schizophrenia	ICD-10: F20-F29	A group of severe mental disorders in which a person has trouble telling the difference between real and unreal experiences, thinking logically, having normal emotional responses to others, and behaving normally in social situations. Symptoms include seeing, hearing, feeling things that are not there, having false ideas about what is taking place or who one is, nonsense speech, unusual behavior, lack of emotion, and social withdrawal.

TABLE A3– COMORBIDITY: BD AND OTHER MENTAL DISORDERS

	BD + Depression	BD + Schizophrenia	BD + Depression + Schizophrenia
All	6277	3093	1164
pre-1956	2236	1114	392
post-1956	4041	1979	772

TABLE A4 – OLS: MENTAL HEALTH DISORDERS AND CAREER OUTCOMES - EXCLUDING YEARS 2013-2015

	ln(Earnings)		P(Earnings = 0)		P(Disability)	
	(1)	(2)	(3)	(4)	(5)	(6)
BD	-0.452***	-0.407***	0.139***	0.117***	0.117***	0.089***
	(0.009)	(0.012)	(0.002)	(0.003)	(0.002)	(0.002)
Depression	-0.404***	-0.336***	0.137***	0.090***	0.061***	0.034***
	(0.003)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)
Schizophrenia	-1.322***	-1.295***	0.440***	0.380***	0.392***	0.380***
	(0.011)	(0.015)	(0.002)	(0.003)	(0.002)	(0.002)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	No	Yes	No	Yes	No	Yes
Mean of Dep. Var.	--	--	.134	.102	.059	.047
R-squared	0.048	0.323	0.046	0. 344	0.089	0.424
N	36,928,892	27,271,306	42,559,551	30,368,570	42,559,551	35,077,362

Standard errors in parentheses are clustered at the individual level.
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable is the natural logarithm of earnings, defined as the sum of all wages and income from self-employment (columns 1-2); an indicator for individuals receiving zero earnings in a given year (columns 3-4); and for individuals on disability (columns 5-6). Earnings are measured in nominal DKK and are the sum of all wages and income from self-employment. The variables *BD, Depression, Schizophrenia* equal 1 for individuals who have been diagnosed with these conditions at least once between 1995 and 2015. We use diagnoses data for calendar years 1995-2012. All regressions include cohort and year fixed effects; columns 2, 4, and 6 include family fixed effects. The sample is restricted to people aged 20-60 born in cohorts 1946-1977; in column 1, we further restrict the sample to people with positive earnings.

TABLE A5 – OLS: MENTAL HEALTH CONDI DISORDERS, ACCESS TO TREATMENT, AND CAREER OUTCOMES - EXCLUDING YEARS 2013-2015

	ln(Earnings)		P(Earnings=0)		P(Disability)	
	(1)	(2)	(3)	(4)	(5)	(6)
BD	-0.569***	-0.668***	0.197***	0.193***	0.218***	0.204***
	(0.019)	(0.043)	(0.004)	(0.010)	(0.004)	(0.010)
BD x post	0.161***	0.271***	-0.085***	-0.085***	-0.147***	-0.130***
	(0.021)	(0.045)	(0.005)	(0.011)	(0.005)	(0.010)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	No	Yes	No	Yes	No	Yes
Mean of Dep. Var.	--	--	.134	.102	.059	.042
R-squared	0.048	0.318	0.046	0.344	0.090	0.424
N	36,928,892	27,271,306	42,559,551	30,368,57	48,071,128	30,368,570

Standard errors in parentheses are clustered at the individual level.

*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable is the natural logarithm of earnings, defined as the sum of all wages and income from self-employment (columns 1-2); an indicator for individuals receiving zero earnings in a given year (columns 3-4); and for individuals on disability (columns 5-6). The variable *BD* equals 1 for individuals who have been diagnosed with this condition at least once between 1995 and 2015. *Post* equals 1 for individuals who were born after 1956, and turned 20 after lithium, the main treatment for bipolar disorder, became available in Denmark in 1976. Controls include indicators for having received at least one diagnosis of *Depression* and *Schizophrenia*. We use diagnoses data for calendar years 1995-2012. All regressions include cohort and year fixed effects; columns 4-6 include family fixed effects. The sample is restricted to people aged 20-60 born in cohorts 1946-1977; in columns 1-2, we further restrict the sample to people with positive earnings.

TABLE A6 – OLS: MENTAL HEALTH DISORDERS AND CAREER OUTCOMES. SAMPLE OF PEOPLE WITH FAMILY IDENTIFIERS

	Ln(Earnings)	P(Earnings = 0)	P(Disability)
	(1)	(2)	(3)
BD	-0.470***	0.145***	0.109***
	(0.010)	(0.003)	(0.002)
Depression	-0.439***	0.137***	0.063***
	(0.004)	(0.001)	(0.001)
Schizophrenia	-1.376***	0.450***	0.428***
	(0.012)	(0.002)	(0.002)
Cohort FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Family FE	No	No	No
Mean of Dep. Var.	--	0.105	0.047
N	31,404,955	35,077,362	35,077,362

Standard errors in parentheses are clustered at the individual level.
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable is the natural logarithm of earnings, defined as the sum of all wages and income from self-employment (column 1); an indicator for individuals receiving zero earnings in a given year (column 2); and for individuals on disability (column 3). Earnings are measured in nominal DKK and are the sum of all wages and income from self-employment. The variables *BD, Depression, Schizophrenia* equal 1 for individuals who have been diagnosed with these conditions at least once between 1995 and 2015. Diagnoses data are available for calendar years 1995-2015. All regressions include cohort and year fixed effects; columns 2, 4, 6, 8, and 10 include family fixed effects. The sample is restricted to people aged 20-60 born in cohorts 1946-1977 for whom family identifiers are available; in column 1, we further restrict the sample to people with positive earnings.

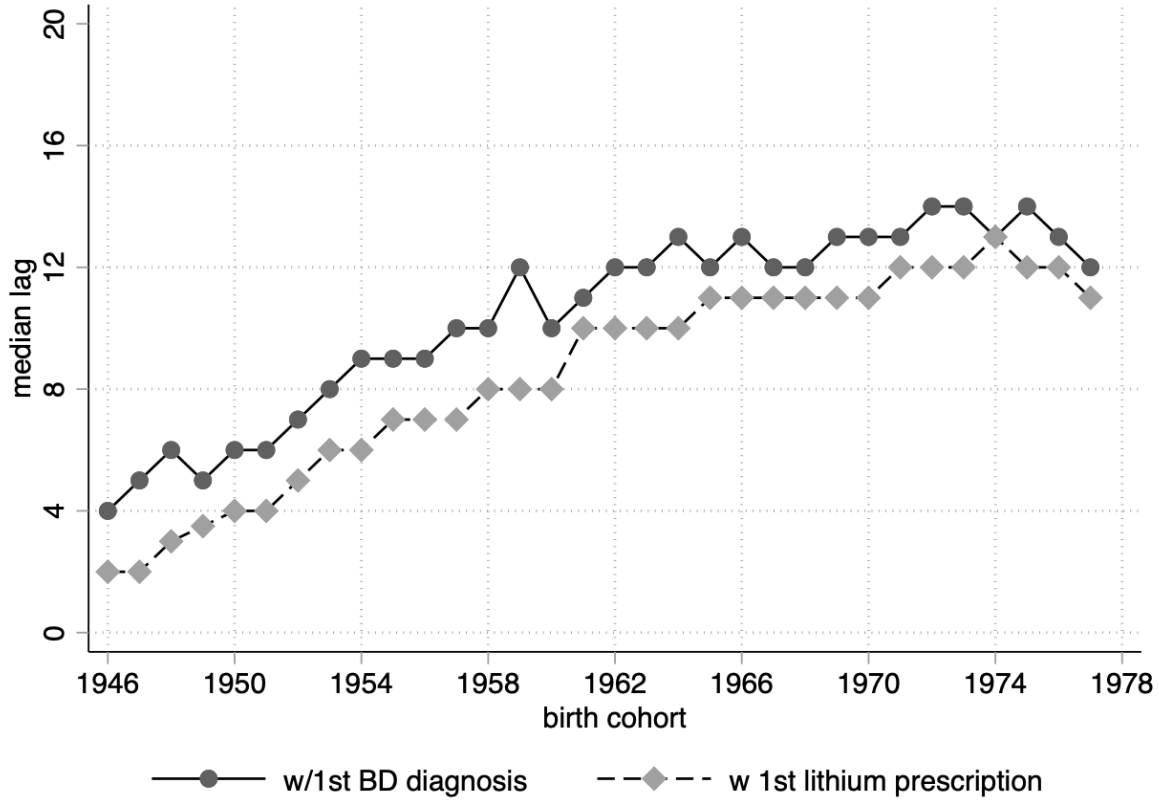
TABLE A7 - OLS: MENTAL HEALTH CONDI DISORDERS, ACCESS TO TREATMENT, AND CAREER OUTCOMES. SAMPLE OF PEOPLE WITH FAMILY IDENTIFIERS

	Ln(Earnings)	P(Earnings = 0)	P(Disability)
	(1)	(2)	(3)
BD	-0.564***	0.188***	0.209***
	(0.031)	(0.007)	(0.007)
BD x post	0.108***	-0.050***	-0.117***
	(0.032)	(0.008)	(0.007)
Cohort FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Family FE	No	Yes	No
Mean of Dep. Var.	--	0.105	0.047
N	31,404,955	35,077,362	35,077,362

Standard errors in parentheses are clustered at the individual level.
*** p<0.01, ** p<0.05, * p<0.1

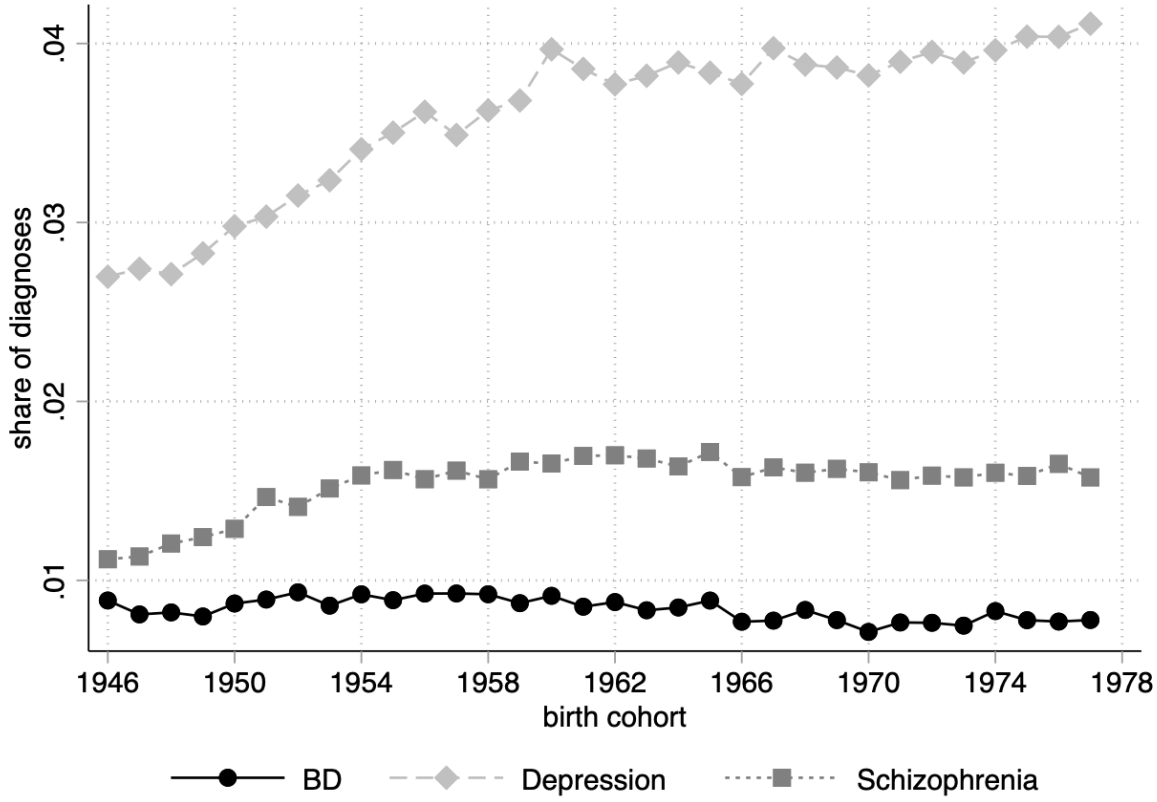
Note: The dependent variable is the natural logarithm of earnings, defined as the sum of all wages and income from self-employment (column 1); an indicator for individuals receiving zero earnings in a given year (column 2); and for individuals on disability (column 3). The variable *BD* equals 1 for individuals who have been diagnosed with this condition at least once between 1995 and 2015. *Post* equals 1 for individuals who were born after 1956, and turned 20 after lithium, the main treatment for bipolar disorder, became available in Denmark in 1976. Controls include indicators for having received at least one diagnosis of *Depression* and *Schizophrenia*. Diagnoses data are available for calendar years 1995-2015. All regressions include cohort and year fixed effects; columns 4-6 include family fixed effects. The sample is restricted to people aged 20-60 born in cohorts 1946-1977 for whom family identifiers are available; in column 1, we further restrict the sample to people with positive earnings.

FIGURE A1– LAG BETWEEN FIRST YEAR IN THE SAMPLE AND FIRST BD DIAGNOSIS/LITHIUM PRESCRIPTION



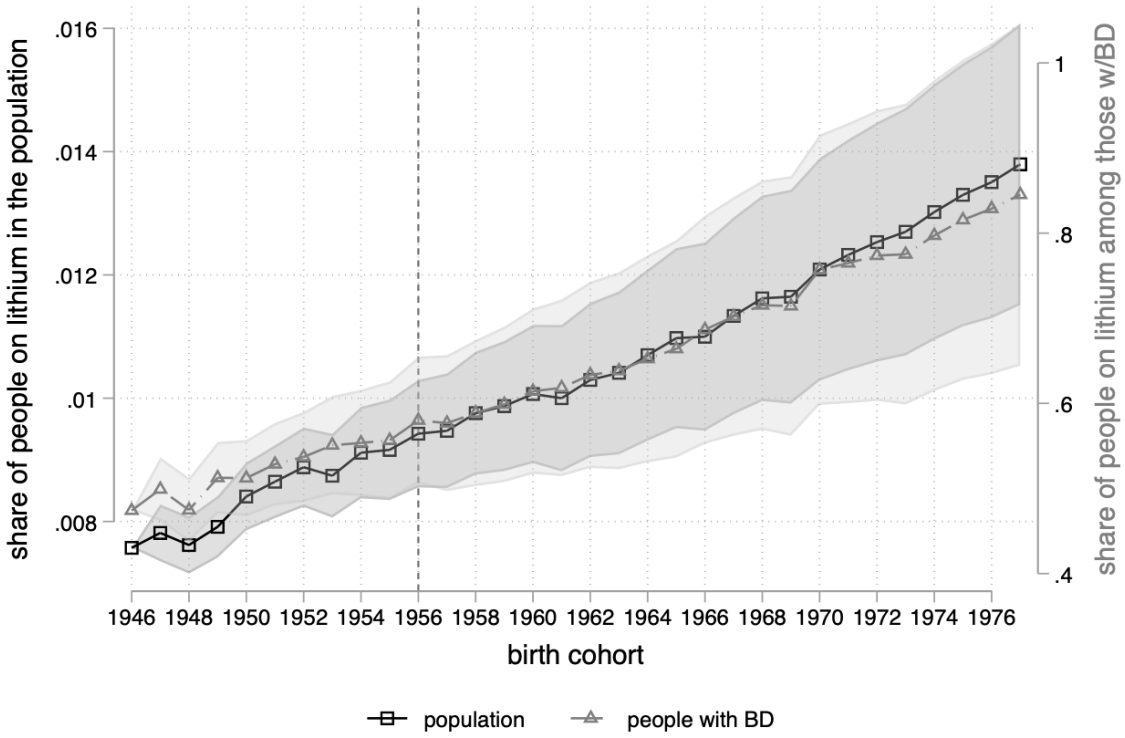
Note: Average lag between the year in which a person entered the sample between 1995 and 2015, and the year in which they received the first diagnosis of BD (solid line) or the first lithium prescription (dashed line).

FIGURE A2– COUNTS OF INDIVIDUALS DIAGNOSED, BY CONDITION AND ACROSS COHORTS



Note: Counts of individuals with at least one diagnosis of depression, BD, or schizophrenia between 1995 and 2015.

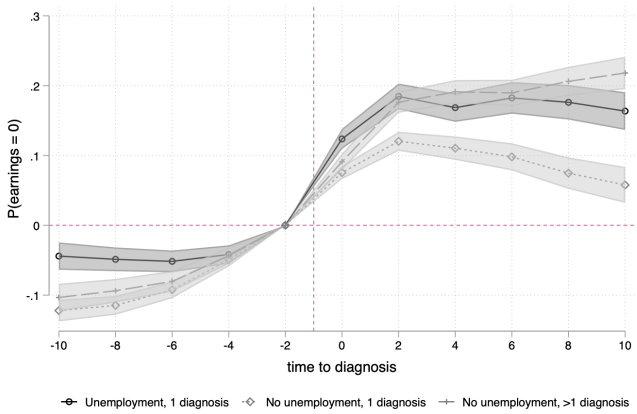
FIGURE A3 – RATES OF LITHIUM PRESCRIPTION, BY COHORT



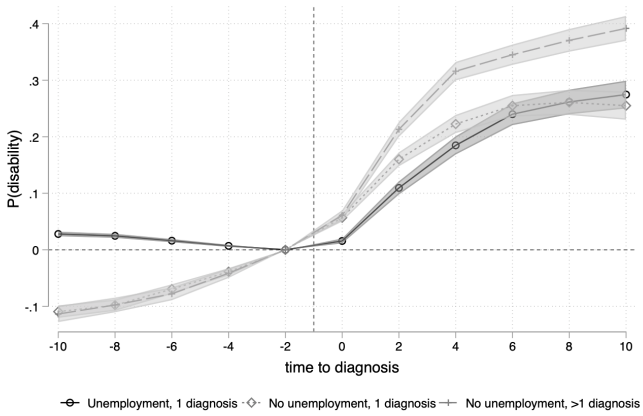
Note: OLS estimates and 95 percent confidence intervals of the parameters δ (cohort fixed effects) in the equation $L_{it} = \alpha_1 a_{it} + \alpha_1 a_{it}^2 + \alpha_1 a_{it}^3 + \delta_{c(i)} + \tau_t + \varepsilon_{it}$, where L_{it} equals one if person i has a lithium diagnosis in year t, a_{it} denotes age, and τ_t are year fixed effects. The darker series (squares) is estimated on the population. The lighter series (triangles) is estimated on people with a BD diagnosis. Standard errors are clustered at the person level.

FIGURE A4– EVENT STUDY OF P(EARNINGS = 0) AND P(DISABILITY) AROUND A DIAGNOSIS BY UNEMPLOYMENT AND #DIAGNOSES

P(EARNINGS = 0)

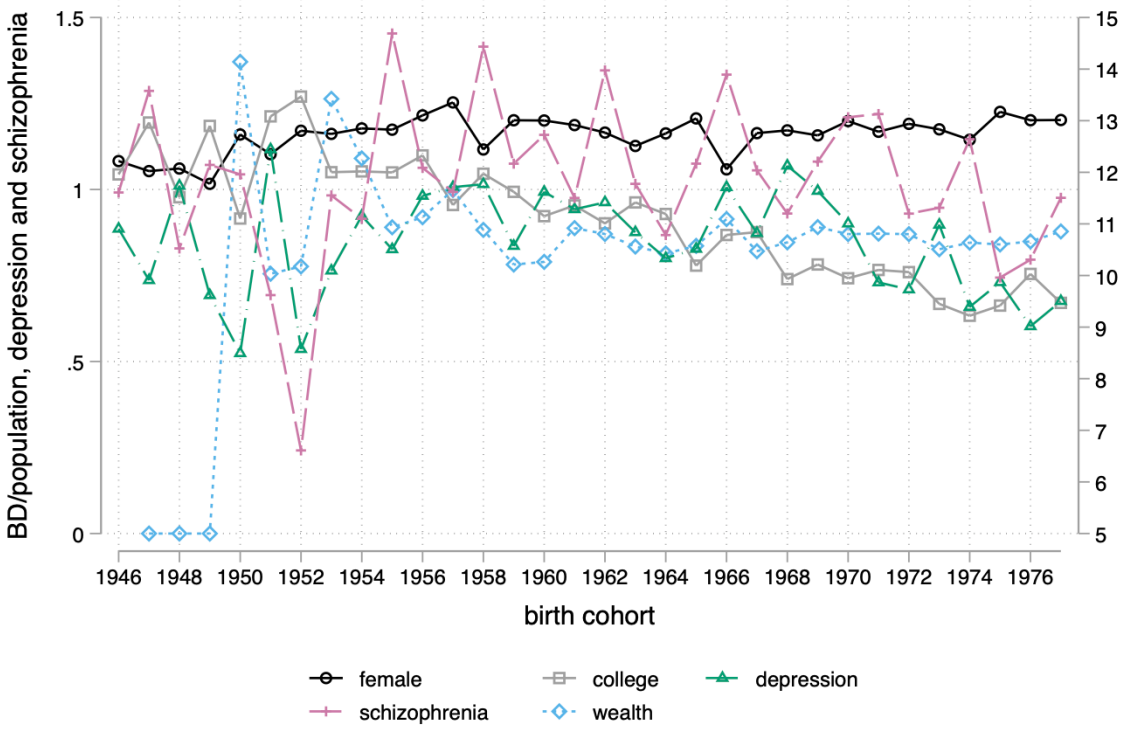


P(DISABILITY)



Note: Point estimates and 95 percent confidence of the parameter δ in equation $Y_{ict} = \sum^{10}_{k=-10} \delta_k C_i I(t\text{-}Y(C)_i = k) + \beta_1 BD_i + \beta_2 Depression_i + \beta_3 Schizophrenia_i + \theta_c + \tau_t + \varepsilon_{ict}$, where the dependent variable is an indicator for having no earnings (top panel) or being on disability (bottom panel), C_i is an indicator for any of *BD*, *Depression*, *or Schizophrenia, $Y(C)_i$* indicates the year when individual *i* is diagnosed with any of these conditions, and *I()* is an indicator function. The vector θ_c contains cohort fixed effects and τ_t are year fixed effects. Standard errors are clustered at the individual level. In the *Unemployment, 1 diagnosis* series, we compare healthy individuals with people with mental health conditions who receive only one diagnosis between 1995 and 2015 and experience at least one unemployment episode in the two years preceding and following the diagnosis. In the *No unemployment, 1 diagnosis* series, we compare healthy individuals with people with mental health conditions who receive only one diagnosis between 1995 and 2015 and do not experience any unemployment in the two years preceding and following the diagnosis. In the *No unemployment, >1 diagnosis* series, we compare healthy individuals with people with mental health conditions who receive more than one diagnosis between 1995 and 2015 and do not experience any unemployment in the two years preceding and following the diagnosis. The sample is restricted to individuals between 20 and 60 years of age, born between 1946 and 1977.

FIGURE A5– CHARACTERISTICS OF PEOPLE WITH BD, BY COHORT

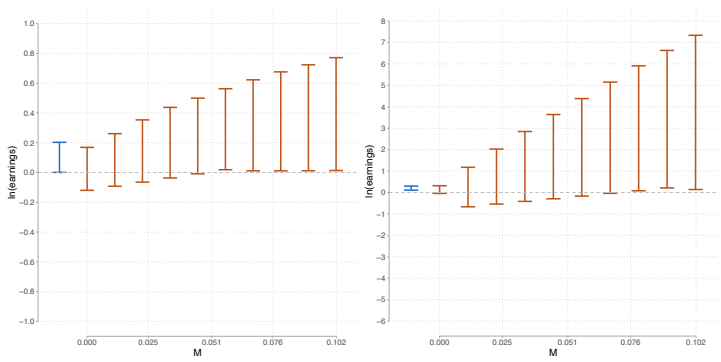


Note: The figure shows ratios of the mean of each variable among people with BD and the mean among the population, separately for each cohort. *Female* equals one for women. *College* equals one for people with at least one college degree. *Depression* and *Schizophrenia* equal one for people with at least one diagnosis of each of these conditions. *Wealth* represents a person's parental wealth. The right vertical axis shows the scale for the *Depression* variable; the left vertical axis shows the scale for all other variables.

FIGURE A6– SENSITIVITY ANALYSIS TO PARALLEL TRENDS ASSUMPTION VIOLATION IN COHORT STUDIES, USING RAMBACHAN AND ROTH (2020) APPROACH

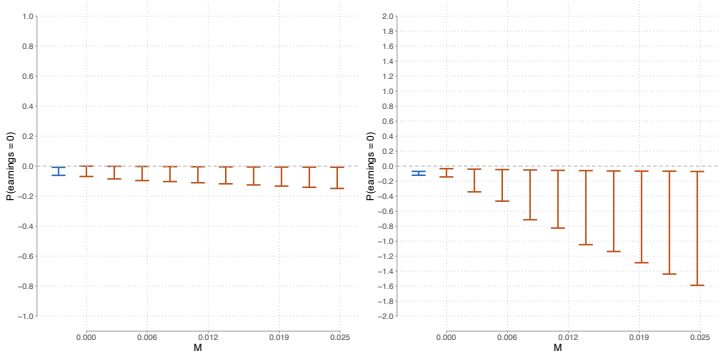
LN(EARNINGS)

1960 1976



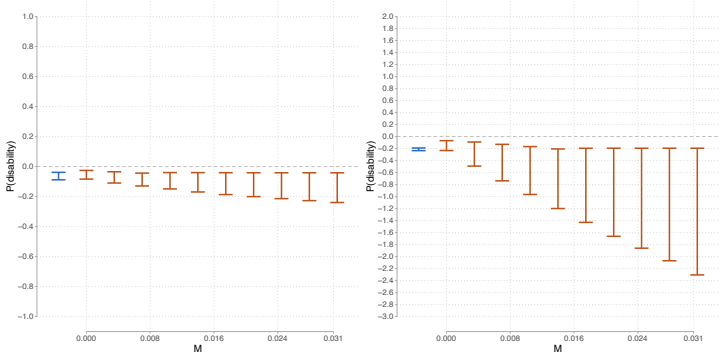
P(EARNINGS=0)

1960 1976



P(DISABILITY=0)

1960 1976



Note: Sensitivity plots of the confidence intervals of $\delta_{1959\text{-}60}$ and $\delta_{1975\text{-}76}$ in equation (5), constructed following the approach of Rambachan and Roth (2020). The blue series represents OLS confidence intervals.