REPLICATION FILE. Code and Data for Evidence and Figures in Brady, David. "Poverty, Not the Poor." *Science Advances*.

FIGURE 1

See Supplementary References at end of References List.

PSID ANALYSES (in text)

This analyses use Brady and Kohler's *WZB-PSID File*, which builds on the PSID and CNEF. See the Stata do file for the WZB-PSID: https://bradydave.files.wordpress.com/2022/07/wzbpsid_2019v4_retrival.pdf. See the codebook for the WZB-PSID File: https://bradydave.files.wordpress.com/2022/07/wzbpsid_2019v4_codebook-1.pdf.

The Dataverse and OSF Project pages contain a data extract including the relevant variables called "WZBPSIDSpells.dta". This is because the PSID requires users of the PSID to register and does not allow posting of data files outside the PSID.

The analyses reported in the text use the following Stata code:

```
*select only relevant variables
keep x11101ll wave inc04eq wght01

* create psid-specific relative poverty measure
gen           relpov = 0
label var     relpov "HH is in relative povery"
label define  relpov 0 "No" 1 "Is poor" .a "Latino sample" .b "Immigrant
sample"
label values  relpov relpov
numlabel      relpov , add mask(#_)
notes         relpov : `tag'
notes         relpov : below are year-specific thresholds for HH poverty, ///

* percent of relative poverty
gen  pctrelpov = .
label var pctrelpov "Percent of relative poverty line"

** loop through years, make relative poverty measurements
qui levelsof wave, local(yrloop)
foreach y in `yrloop' {
  qui sum inc04eq ///
    if wave == `y' [aw=wght01] , det
  local line = r(p50) / 2
  local linenote = round(`line',.001)
  replace relpov = 1 ///
    if inc04eq < `line' & wave == `y'
  notes relpov : Year `y' - `linenote'
  * percent relative poverty
  replace pctrelpov = (inc04eq / r(p50)) * 100 ///
    if wave == `y'
}

tabstat relpov [aw=wght01], by(wave)
```

```
*SPELLS RESULTS
*gen alt wave variable
recode wave (1971=1) (1973=2) (1975=3) (1977=4) (1979=5) (1981=6) (1983=7)
(1985=8) (1987=9) (1989=10) (1991=11) (1993=12) (1995=13) (1997=14) (1999=15)
(2001=16) (2003=17) (2005=18) (2007=19) (2009=20) (2011=21) (2013=22)
(2015=23) (2017=24) (2019=25) (else=.), gen (wavecount)
*Retain odd-years sample
gen odd=0 if wave !=.
replace odd=1 ///
  if inlist(wave, 1971, 1973, 1975, 1977, 1979, 1981, 1983, 1985, 1987, 1989,
1991, 1993, 1995, 1997, 1999, 2001, 2003, 2005, 2007, 2009, 2011, 2013, 2015,
2017, 2019)
keep if odd==1

*tsset panel data
tsset x1110111 wavecount
drop if wave<1971

*count consecutive years observed for each respondent
sort x1110111 wavecount
gen validincome=0 if inc04eq==.
replace validincome=1 if inc04eq!=.
tsspell validincome, seq(consecinc)
drop _spell
drop _end

*estimate spells for relpov
sort x1110111 wavecount
tsspell relpov, seq(lengthr50) spell(numspellr50) end(endr50)
label var lengthr50 "Length of Spell in Years"
label var numspellr50 "Number of Spells"
label var endr50 "Exit Poverty"

*top code length of poverty spells at 20 years
replace lengthr50=10 if lengthr50>10

*label values of lengthr50
label define lengthr50 1 "1-2" 2 "3-4" 3 "5-6" 4 "7-8" 5 "9-10" 6 "11-12" 7
"13-14" 8 "15-16" 9 "17-18" 10 "19-20"

*NOT remove exits because of death

*How long are poverty spells?
tab lengthr50 if relpov & wave==2019 [aw=wght01]
tab lengthr50 if relpov [aw=wght01]

*Lifetime probability of poverty
egen everpoor = max(relpov) if relpov!=., by(x1110111)
tabstat everpoor if wave>2010 & wave<2020 [aw=wght01], stats (mean semean n)
by(wave)
```

****FIGURE 2:**

See Supplementary References at end of References List.

*****MOST REMAINING FIGURES USE AUTHOR'S ANALYSES OF LIS DATA*****

*The LIS is accessible with registration through www.lisdatacenter.org. LIS registration is free for students and residents of member countries (see: http://www.lisdatacenter.org/data-access/access-eligibility/). The LIS data may not be downloaded, and must be accessed remotely through the LISSY interface (see: http://www.lisdatacenter.org/data-access/lissy/). Do files can be submitted to LISSY as "jobs", and results are returned as text via the interface and email.

*Analyses conducted: March, 2023
*Analyses updated June, 2023 to include recent Swedish and U.S. datasets and for replication

BUILD DATAFILE USING U.S. DATASETS 2016-2020 and 30 Peer Rich Democracies Most Recent Wave Through 2019

```
global c "us16 us17 us18 us19 us20 au18 at19 be17 ca18 cz16 dk16 ee16 fi16 fr18 de19 gr16 hu15 is10 ie19 il18 it16 jp13 lt18 lu19 nl18 no19 pl19 sk18 si15 kr16 es16 se19 ch18 uk19 tw16"

foreach x of global c {

*HH file
use $`x'h, clear

*Drop missing and zero dhi
drop if dhi==.
drop if dhi==0
drop if hwgt==.

*Equivalize disposable HH income as square root of household member
generate ey=(dhi/(nhhmem^0.5))

* create person weight as hwgt times number of household member
generate wt=hwgt*nhhmem

*Top- and bottom-code EY
quietly sum ey [w=wt]
generate botlin=0.01*_result(3)
replace ey=botlin if ey<botlin
quietly sum ey [w=wt], de
generate toplin=10*_result(10)
replace ey=(toplin/(nhhmem^0.5)) if ey>toplin

*Poverty threshold as <50% of median EY
quietly sum ey [w=wt], de
quietly generate povl=_result(10)*.5

*Define POOR
gen poor=.
replace poor=0 if ey>=povl & ey!=.
replace poor=1 if ey<povl

*Other HH-level variables
gen multearn=.
replace multearn=0 if nearn==0 | nearn==1
```

```
replace multearn=1 if nearn>1 & nearn!=.
gen unemphh=.
replace unemphh=0 if nearn>0 & nearn!=.
replace unemphh=1 if nearn==0
gen emphh=.
replace emphh=0 if nearn==0
replace emphh=1 if nearn>0 & nearn!=.

*geographic variables

sort hid
keep did year hid hwgt iso2 poor povl nhhmem hhtype nhhmem65 nhhmem17 hi* dhi
hpartner nearn unemphh multearn emphh dname cname locsz_c area_c region_c

save $mydata/`x'hSA, replace

*Person File
use $`x'p, clear

*Sex
recode sex (1=0)(2=1)(.=.), gen(female)
recode sex (1=1)(2=0)(.=.), gen(male)

*Identifying heads
gen head=.
replace head=1 if relation==1000
replace head=0 if relation>1000 & relation!=.
sort hid
keep hid pid did year relation partner nchildren ageyoch parents age sex
ethnic_c immigr citizen ctrybrth yrsresid educ emp pi* pilabour gross1 educ_c
marital head male female

save $mydata/`x'pSA, replace

merge m:1 hid using $mydata/`x'hSA, keep(match) nogen

*create variable for lead earner*
recode pilabour (.=0)
egen maxinc=max(pilabour), by(did hid)
gen lead=pilabour==maxinc
egen maxage=max(age) if lead, by(did hid)
replace lead=0 if age~=maxage
egen numlead = sum(lead), by(did hid)
gen rlead = runiform()
egen maxrlead = max(rlead) if lead, by(did hid)
replace lead = 0 if numlead>1 & rlead<maxrlead

*create variables for education*
gen leadeduc_a=educ*lead
egen leadeduc=max(leadeduc_a), by(hid)
recode leadeduc (3=1) (nonmiss=0), gen(highed)
recode leadeduc (1=1)(nonmiss=0), gen(lowed)
gen agelead_a=age*lead
egen agelead=max(agelead_a), by(hid)
gen ageleadsq=agelead^2

*create family structure variables*
```

```
gen married=.
replace married=0 if marital>=200 & marital!=.
replace married=1 if marital<200 | partner==110
gen marriedhh_a=married*head
egen marriedhh=max(marriedhh_a), by(hid)
recode marriedhh (1=0)(0=1)(.=.), gen(single)

*singmom (based on nchildren & ageyoch)
recode nchildren 2/17=1, gen(nchild)
replace nchild=0 if ageyoch>17 & ageyoch!=.
gen sing_mom_a=head*female
gen sing_mom_b=sing_mom_a*single
gen sing_mom_c=sing_mom_b*nchild
replace sing_mom_c=0 if age>54
egen singmom=max(sing_mom_c), by(hid)
replace singmom=1 if singmom>1 & singmom!=.
gen sing_dad_a=head*male
gen sing_dad_b=sing_dad_a*single
gen sing_dad_c=sing_dad_b*nchild
egen singdad =max(sing_dad_c), by(hid)
replace singdad=1 if singdad>1 & singdad!=.
gen fhnk_a=0
replace fhnk_a=1 if sing_mom_b ==1 & nhhmem17==0
egen fhnk=max(fhnk_a), by(hid)
gen mhnk_a=0
replace mhnk_a=1 if sing_dad_b ==1 & nhhmem17==0
egen mhnk=max(mhnk_a), by(hid)

*lead age groups
gen leadu25=0 if agelead!=.
replace leadu25=1 if agelead<25 & agelead~=.
gen lead2534=0 if agelead!=.
replace lead2534=1 if agelead>24 & agelead<35
gen lead3554=0 if agelead!=.
replace lead3554=1 if agelead>34 & agelead<55
gen leado54=0 if agelead!=.
replace leado54=1 if agelead>54 & agelead~=.

*lead immigration
gen leadimmig_a=lead*immigr
egen leadimmig=max(leadimmig_a), by(hid)
gen noncit=0 if citizen==1000 | citizen==1300
replace noncit=1 if citizen==2000
gen leadnoncit_a=lead*noncit
egen leadnoncit=max(leadnoncit_a), by(hid)
save $mydata/`x'SA, replace
}

global d " us17SA us18SA us19SA us20SA au18SA at19SA be17SA ca18SA cz16SA
dk16SA ee16SA fi16SA fr18SA de19SA gr16SA hu15SA is10SA ie19SA il18SA it16SA
jp13SA lt18SA lu19SA nl18SA no19SA pl19SA sk18SA si15SA kr16SA es16SA se19SA
ch18SA uk19SA tw16SA"

use $mydata/us16SA, clear
foreach x of global d {
append using "$mydata/`x'"
}
```

```
save $mydata/June23SA, replace
*/
tab dname
tab cname
```

***Poverty for 4 Major Risk Groups (in text)**

```
use $mydata/June23SA, clear

tabstat poor if unemphh==1 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)
tabstat poor if lowed==1 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)
tabstat poor if singmom==1 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)
tabstat poor if leadu25==1 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)
```

*****Figure 3***

```
use $mydata/June23SA, clear

tabstat poor if age<18 [w=hwgt], stats(mean sem n) by(cname)

keep if cname=="United States"

*Race/ethnicity
recode ethnic_c (1 = 1 "White") (2 4 6 8 10 12 = 2 "Latino") (3 = 3 "Black")
( 7 9 = 4 “Asian”) ( 5 = 5 “Native”), gen(race)

*race differences*
tabstat poor if age<18  [w=hwgt], stats(mean sem n) by(race)
```

*****Figure 4***

```
**This figure is generated with a do file as a separate job – not using the
file built above and used in other figures.

global c " us70 us71 us72 us74 us75 us76 us79 us80 us81 us84 us85 us86 us89
us90 us91 us94 us95 us96 us99 us00 us01 us04 us05 us06 us09 us10 us11 us14
us15 us16 us18 us19 us20"
foreach x of global c {

*HH file
use $`x'h, clear

*drop missing
drop if dhi==.
drop if dhi==0
drop if hwgt==.
drop if hwgt==0

*equivalize and top and bottom-code income
gen wt=hwgt*nhhmem
gen ey=dhi/(sqrt(nhhmem))
qui sum ey [w=wt]
gen botlin=0.01*_result(3)
replace ey=botlin if ey<botlin
```

```
quietly sum dhi [w=wt], de
gen toplin=10*_result(10)
replace ey=(toplin/(nhhmem^0.5)) if dhi>toplin

*Poverty threshold
quietly sum ey [w=wt], de
generate povl5=_result(10)*.5

*Define poverty
gen poor5=.
replace poor5=0 if ey>= povl5 & ey!=.
replace poor5=1 if ey< povl5 & ey!=.

sort hid
keep hid did year dname cname hwgt ey dhi poor5

save $mydata/brady/`x'h, replace

*Person File
use $`x'p, clear
recode sex (1=0)(2=1)(.=.), gen(female)
recode sex (1=1)(2=0)(.=.), gen(male)
sort hid
keep hid pid did year age sex ethnic_c male female
save $mydata/brady/`x'p, replace
merge m:1 hid using $mydata/brady/`x'h, keep(match) nogen
save $mydata/brady/`x', replace
}

*** append country files
global c " us71 us74 us75 us76 us79 us80 us81 us84 us85 us86 us89 us90 us91
us94 us95 us96 us99 us00 us01 us04 us05 us06 us09 us10 us11 us14 us15 us16
us18 us19 us20"
use $mydata/brady/us70, clear

foreach x of global c {
append using "$mydata/brady/`x'"
}

tab dname
save $mydata/brady/income, replace

*1965
tabstat poor5 if age<18 & year>1973 & year<1977 [w=hwgt], by(ethnic_c)
stats(mean sem n)

*1970
*race coding available after 1970
recode ethnic_c (11 = 1 "White") (21 = 2 "Black") (12 22 92 = 3 "Latino")
(else=4 "Other") if year>1969, gen(race)
tabstat poor5 if age<18 & year>1969 & year<1973 [w=hwgt], by(race) stats(mean
sem n)

*1975
tabstat poor5 if age<18 & year>1973 & year<1977 [w=hwgt], by(race) stats(mean
sem n)
```

```
*1980
tabstat poor5 if age<18 & year>1978 & year<1982 [w=hwgt], by(race) stats(mean
sem n)

*1985
tabstat poor5 if age<18 & year>1983 & year<1987 [w=hwgt], by(race) stats(mean
sem n)

*1990
*new race coding available
drop race
recode ethnic_c (1 = 1 "White") (2 4 6 8 10 12 = 5 "Latino") (3 = 3 "Black")
(else=4 "Other") if year>1988, gen(race)
tabstat poor5 if age<18 & year>1988 & year<1992 [w=hwgt], by(race) stats(mean
sem n)

*1995
tabstat poor5 if age<18 & year>1993 & year<1997 [w=hwgt], by(race) stats(mean
sem n)

*2000
tabstat poor5 if age<18 & year>1998 & year<2002 [w=hwgt], by(race) stats(mean
sem n)

*2005
tabstat poor5 if age<18 & year>2003 & year<2007 [w=hwgt], by(race) stats(mean
sem n)

*2010
tabstat poor5 if age<18 & year>2008 & year<2011 [w=hwgt], by(race) stats(mean
sem n)

*2015
tabstat poor5 if age<18 & year>2013 & year<2016 [w=hwgt], by(race) stats(mean
sem n)

*2019
tabstat poor5 if age<18 & year>2017 & year<2021 [w=hwgt], by(race) stats(mean
sem n)
```

*****Figure 5***

```
use $mydata/June23SA, clear

*Overall (get US & cross-national median from this)
tabstat poor [w=hwgt], stats(mean sem n) by(cname)

*Working-Aged HHs
tabstat poor if agelead>17 & agelead<65 [w=hwgt], stats(mean sem n) by(cname)

*Employed HHs
tabstat poor if emphh==1 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)

*High School+ HHs
```

```
tabstat poor if lowed==0 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)

*Coupled HHs
tabstat poor if marriedhh==1 & agelead>17 & agelead<65 [w=hwgt], stats(mean
sem n) by(cname)

*Non-Young HHs
tabstat poor if leadu25==0 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)

*Zero Risk HHs
gen norisk=1 if unemphh==0 & lowed==0 & singmom==0 & leadu25==0
replace norisk=0 if unemphh==1| lowed==1| singmom==1| leadu25==1
tabstat poor if norisk==1 & agelead>17 & agelead<65 [w=hwgt], stats(mean sem
n) by(cname)

*population shares in non-risk groups (in text)*
tabstat emphh unemphh if agelead>17 & agelead<65 [w=hwgt], stats(mean sem n)
by(cname)
tabstat lowed highed if agelead>17 & agelead<65 [w=hwgt], stats(mean sem n)
by(cname)
tabstat marriedhh if agelead>17 & agelead<65 [w=hwgt], stats(mean sem n)
by(cname)

*share of US population in working poor or jobless poor HHs (in text)
gen jobless=0 if emphh==1 | agelead<17 | agelead>65
replace jobless=1 if emphh==0 & agelead>17 & agelead<65
gen working=0 if emphh==0 | agelead<17 | agelead>65
replace working=1 if emphh==1 & agelead>17 & agelead<65
gen workpoor=working*poor
gen joblesspoor=jobless*poor
tabstat workpoor joblesspoor [w=hwgt], stats(mean sem n)
```

*****Figure 6***

```
use $mydata/June23SA, clear

keep if cname=="United States"

*Under 18
gen kid=0 if age!=.
replace kid=1 if age<18
tabstat poor if kid==1 [w=hwgt], stats(mean sem n)

*18-64
gen adult=0 if age!=.
replace adult=1 if age>17 & age<65
tabstat poor if adult==1 [w=hwgt], stats(mean sem n)

*>64
gen elder=0 if age!=.
replace elder=1 if age>64 & age!=.
tabstat poor if elder==1 [w=hwgt], stats(mean sem n)

*Female adults
tabstat poor if age>17 & age!=. & female==1 [w=hwgt], stats(mean sem n)
```

```
*Male Adults
tabstat poor if age>17 & age!=. & female==0 [w=hwgt], stats(mean sem n)

*Regions
gen south =0 if region_c!=.
replace south=1 if region_c== 51| region_c== 52| region_c== 53| region_c==
54| region_c== 55| region_c== 56| region_c== 57| region_c== 58| region_c==
59| region_c== 61| region_c== 62| region_c== 63| region_c== 64| region_c==
71| region_c== 72| region_c== 73| region_c== 74
gen midwest =0 if region_c!=.
replace midwest=1 if region_c== 31| region_c== 32| region_c== 33| region_c==
34| region_c== 35| region_c== 41| region_c== 42| region_c== 43| region_c==
44| region_c== 45| region_c== 46| region_c== 47
gen west =0 if region_c!=.
replace west=1 if region_c== 81| region_c== 82| region_c== 83| region_c== 84|
region_c== 85| region_c== 86| region_c== 87| region_c== 88| region_c== 91|
region_c== 92| region_c== 93| region_c== 94| region_c== 95
gen northeast =0 if region_c!=.
replace northeast =1 if region_c== 11| region_c== 12| region_c== 13|
region_c== 14| region_c== 15| region_c== 16| region_c== 21| region_c== 22|
region_c== 23

tabstat poor if south==1 [w=hwgt], stats(mean sem n)
tabstat poor if northeast==1 [w=hwgt], stats(mean sem n)
tabstat poor if midwest==1 [w=hwgt], stats(mean sem n)
tabstat poor if west==1 [w=hwgt], stats(mean sem n)

*Urban/rural ares
*MSA defined as residing in MSA
gen msa=0 if area_c!=.
replace msa=1 if area_c==1 | area_c==2

*metro defined as >=100k
gen metro=0 if locsz_c !=.
replace metro=1 if locsz_c>1 & locsz_c!=.

*Central defined as central city (not "balance of msa")
gen central=0 if area_c!=.
replace central=1 if area_c==1

tabstat poor if msa==1 [w=hwgt], stats(mean sem n)
tabstat poor if metro==1 [w=hwgt], stats(mean sem n)
tabstat poor if central==1 [w=hwgt], stats(mean sem n)

gen suburb=0 if metro!=.
replace suburb=1 if metro==1 & central==0
gen rural=0 if metro!=.
replace rural=1 if metro==0

tabstat poor if suburb==1 [w=hwgt], stats(mean sem n)
tabstat poor if rural==1 [w=hwgt], stats(mean sem n)
```

*****Figure 7***

```
use $mydata/June23SA, clear
```

```
keep if cname=="United States"

*single moms, young single mom and coupled hhs
gen workagecouplehh=0 if marriedhh==0 | agelead<17 | agelead>65
replace workagecouplehh=1 if marriedhh==1 & agelead>17 & agelead<65
gen ysingmom=0 if singmom==0 | leadu25==0
replace ysingmom=1 if singmom==1 & leadu25==1

tabstat singmom ysingmom workagecouplehh if poor==1 [w=hwgt], stats(mean sem
n) by(cname)

*jobless vs employed hhs
gen jobless=0 if emphh==1 | agelead<17 | agelead>65
replace jobless=1 if emphh==0 & agelead>17 & agelead<65
gen working=0 if emphh==0 | agelead<17 | agelead>65
replace working=1 if emphh==1 & agelead>17 & agelead<65

tabstat working jobless if poor==1 [w=hwgt], stats(mean sem n) by(cname)

*high school dropouts vs at least high school
gen nonlowed=0 if lowed==1 | agelead<17 | agelead>65
replace nonlowed=1 if lowed==0 & agelead>17 & agelead<65

tabstat lowed nonlowed if poor==1 [w=hwgt], stats(mean sem n) by(cname)

*Black and Latino adolescents
gen adoles=0 if age!=.
replace adoles=1 if age>9 & age<20
gen black=0 if ethnic_c!=.
replace black=1 if ethnic_c==3
gen blackado=0 if black!=. & adoles!=.
replace blackado=1 if black==1 & adoles==1
gen latino=0 if ethnic_c!=.
replace latino=1 if ethnic_c==2| ethnic_c== 4| ethnic_c== 6| ethnic_c== 8|
ethnic_c==10| ethnic_c==12
gen latinado=0 if latino!=. & adoles!=.
replace latinado=1 if latino==1 & adoles==1
gen central=0 if area_c!=.
replace central=1 if area_c==1
gen blackadocent=0 if blackado==0 | central==0
replace blackadocent=1 if blackado==1 & central==1
gen latinadocent=0 if latinado==0 | central==0
replace latinadocent=1 if latinado==1 & central==1

tabstat blackado blackadocent latinado latinadocent if poor==1 [w=hwgt],
stats(mean sem n) by(cname)

*few or multiple risks
gen sumrisk= unemphh + lowed + singmom + leadu25
gen lowrisk= 0 if sumrisk>1 & sumrisk!=.
replace lowrisk=1 if sumrisk<2
gen highrisk= 0 if sumrisk<2
replace highrisk=1 if sumrisk>1 & sumrisk!=.

tabstat lowrisk highrisk if poor==1 [w=hwgt], stats(mean sem n)

**HOMELESSNESS - SEE SOURCE IN TEXT
```

*****Figure 8: Sources***

```
*For poverty rates: See LIS Key Figures in Supplementary References at end of
References List.
*For social welfare expenditures as % of GDP: www.stats.oecd.org
```

The Dataverse and OSF Project pages contain a data file for these variables titled: "SocPoverty_Final".

```
*Data imported into Stata v17
reghdfe povertyrate socx socxsq if miss==0, absorb(idn year) vce(cluster idn)
margins, at(socx=(9(5)33)) vsquish
marginsplot, xlabel(9(5)33) ytitle(Pr(Poverty Rate)) addplot(scatter
povertyrate socx if miss==1, msymbol(none) mlabel(id) mlabposition(0))
legend(off)
```

*****Figure 9***

```
use $mydata/recentSA, replace

tabstat poor if leadimmig==1 [w=hwgt], stats(mean sem n) by(cname)
tabstat poor if leadimmig==0 [w=hwgt], stats(mean sem n) by(cname)
tabstat poor if leadimmig==1 & leadnoncit==1 [w=hwgt], stats(mean sem n)
by(cname)
tabstat poor if leadimmig==1 & leadnoncit==0 [w=hwgt], stats(mean sem n)
by(cname)
```