



Multiple Imputation with Massive Data: an Application to the Panel Study of Income Dynamics

July 19, 2020

Yajuan Si^{1*}

Steve Heeringa¹

David Johnson¹

Roderick Little²

Wenshuo Liu³

Fabian Pfeffer¹

Raghunathan Trivellore²

1. Survey Research Center, University of Michigan, Ann Arbor, MI, USA; 2. Department of Biostatistics and Survey Research Center, University of Michigan, Ann Arbor, MI, USA; 3. Interactions LLC.

Acknowledgements: Funding for this research was provided by the Michigan Center on the Demography of Aging (P30-AG012846), the National Institute on Aging (R01-AG040213), the National Science Foundation (SES 1623864), and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (R01-HD069609 and P01-HD087155).

*Corresponding email: yajuan@umich.edu

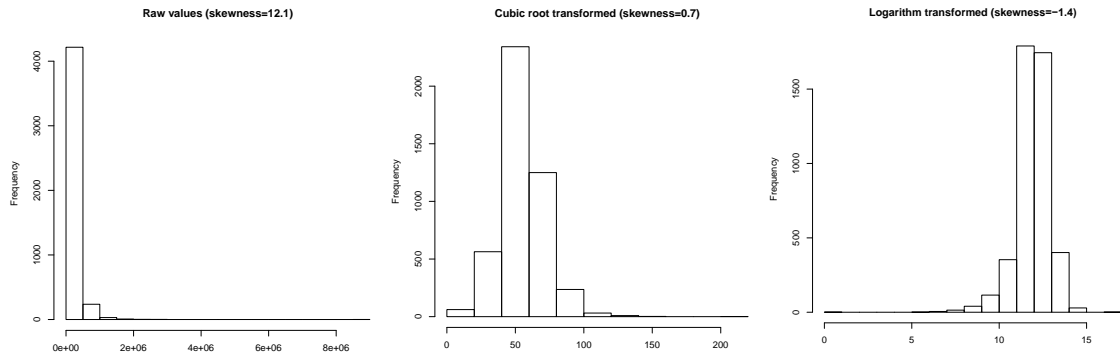


Figure 2: The distribution and skewness of home values before and after transformation.

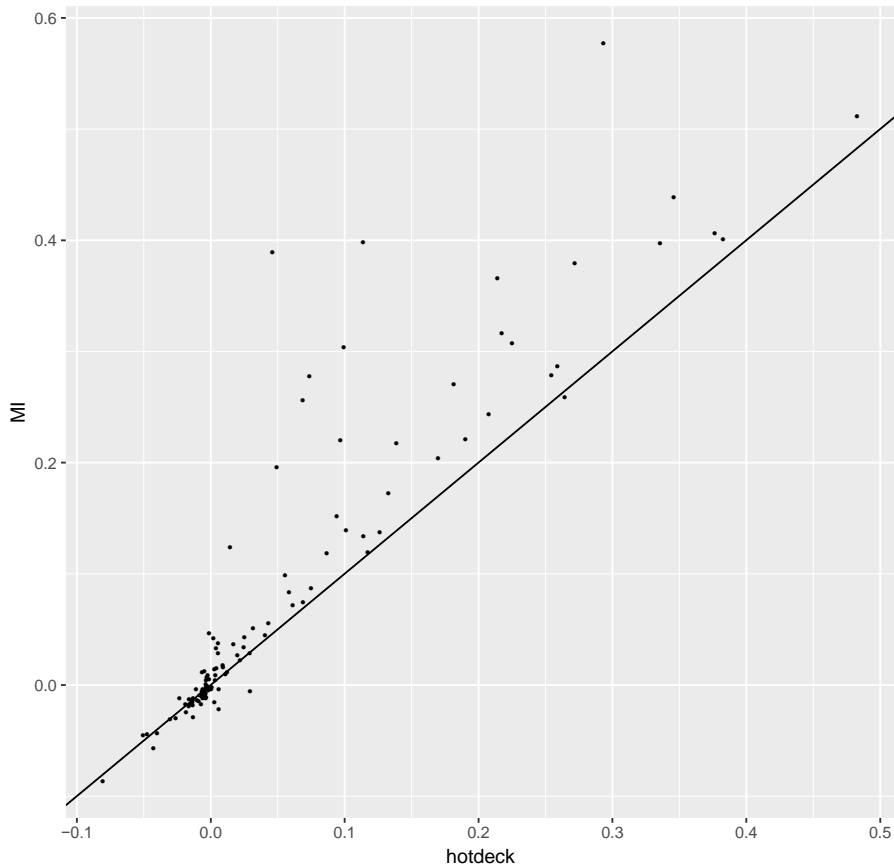


Figure 3: Pairwise correlation coefficients between 16 wealth components.

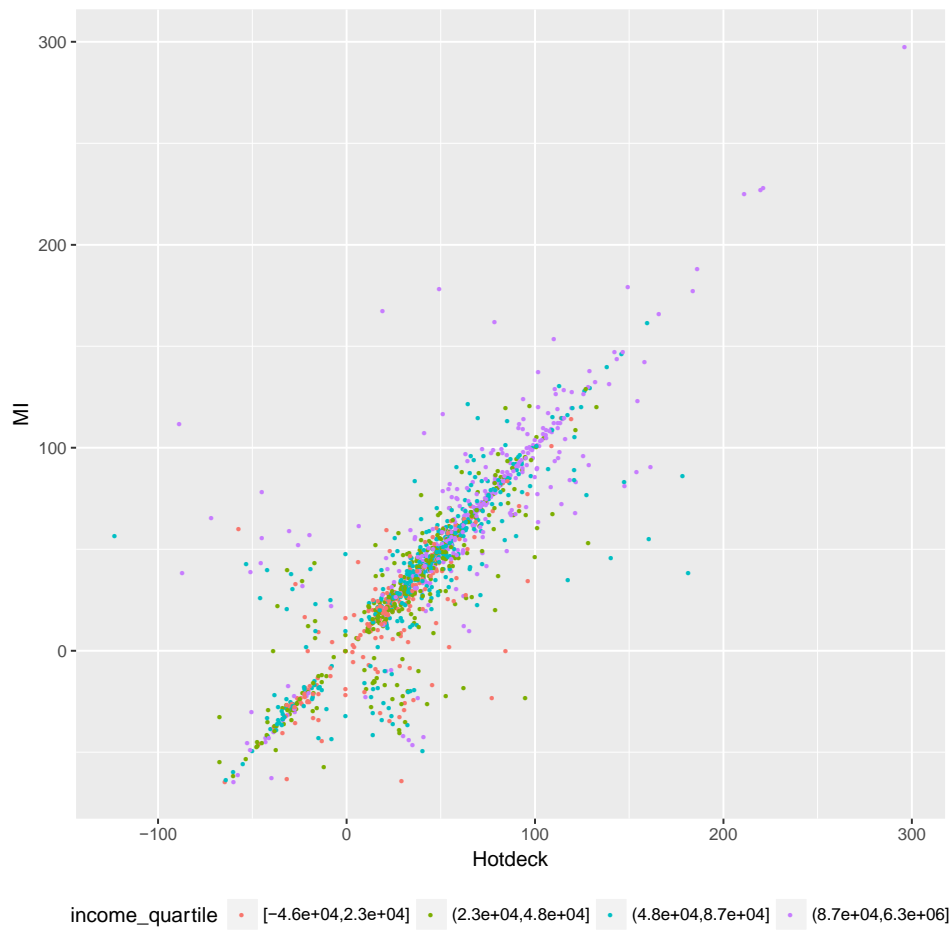


Figure 4: Comparison of imputed wealth values (cubic root transformed) from MI and hot deck imputation. The observations are marked in color grouped by income quartiles (in thousands): [-46, 23], (23, 48], (48, 87], and (87, 630]. Some wealth values are negative because of high debts and low assets.

Appendix 1: Names of variables included for wealth imputation in the 2013

PSID data (The labels are available in <https://simba.isr.umich.edu/VS/s.aspx>)

"AQCASE" "A4" "A8" "A19" "A20" "A21" "A22" "A23" "A23A_1" "A23B_1" "A24_1" "A28_1" "A23A_2"
"A23B_2" "A24_2" "IMP_A20" "A20L" "A20H" "IMP_A19" "IMP_A23" "IMP_A24_1" "IMP_A28_1"
"IMP_A24_2" "A24_1L" "A24_1H" "A24_2L" "A24_2H" "W1" "W1A" "W2A" "W2B" "W6" "W10" "W11A"
"W11B" "W15" "W16" "W21" "W21A" "W22" "W27" "W28" "W33" "W34" "W38A" "W38BLEGL" "W38BMED"
"W38BOTR" "W38BRELS" "W38BSTU" "W39A" "W39B1" "W39B2" "W39B3" "W39B4" "W39B7"
"IMP_W11A" "W11AL" "W11AH" "IMP_W11B" "W11BL" "W11BH" "IMP_W28" "W28L" "W28H"
"IMP_W2A" "W2AL" "W2AH" "IMP_W2B" "W2BL" "W2BH" "IMP_W16" "W16L" "W16H" "IMP_W6" "W6L"
"W6H" "IMP_W34" "W34L" "W34H" "IMP_W22" "W22L" "W22H" "IMP_W39A" "W39AL" "W39AH"
"IMP_W39B1" "W39B1L" "W39B1H" "IMP_W39B2" "W39B2L" "W39B2H" "IMP_W39B3" "W39B3L"
"W39B3H" "IMP_W39B4" "W39B4L" "W39B4H" "IMP_W39B7" "W39B7L" "W39B7H" "IMP_W1" "IMP_W10"
"IMP_W15" "IMP_W21" "IMP_W27" "IMP_W33" "IMP_W38A" "IMP_W38BSTU" "IMP_W38BMED"
"IMP_W38BLEGL" "IMP_W38BRELS" "IMP_W38BOTR" "KEY_MISS" "NFU" "NKIDS" "NELDERS"
"NFUHUCOLLEGE" "NOFUMCOLLEGE" "H49" "D1CKPT" "TOTHU" "GENMARSTAT" "IWLENMINS"
"CCS" "G5" "G99" "G100" "G101" "G103" "G104" "G106" "G107" "G108" "G109" "G110" "G112" "G113"
"G114" "WTRMETRO" "BNUM" "HBNUM" "WBNUM" "HWBIZ" "FARMY" "HDWG" "BNUS" "OVTM" "TIP"
"COMS" "OTHINC" "PROF" "XTRA" "HDRENT" "HDDIV" "HDINT" "HDTRUST" "WFEARN" "WFRENT"
"WFDIV" "WFINT" "WFTRUST" "TAXHW" "HDADC" "HDSSI" "HDOTRWELF" "HDVA" "HDPENS"
"HDANN" "HDIRA" "HDUNEMP" "HDWRKCOMP" "HDCHSUP" "HDALIMONY" "HDHELPREL"
"HDHELPPNON" "HDMISC" "WFADC" "WFSSI" "WFOTRWELF" "WFPENS" "WFANN" "WFIRA"
"WFUNEMP" "WFWRKCOMP" "WFCHSUP" "WFHELPREL" "WFHELPPNON" "WFMISC" "OFLAB"
"OFASSET" "OFADC" "OFSSI" "OFOTRWELF" "OFVA" "OPENS" "OFUNEMP" "OFCHSUP" "OFMISC"
"HDSSEC" "WFSSEC" "OFSSEC" "TOTFAMY" "FDHM13" "FDOUT13" "FDDEL13" "RENT13" "UTIL13"
"TELINT13" "VEHLN13" "VEHPAY13" "VEHLS13" "AUTOIN13" "VEHADD13" "VEHREP13" "GAS13"
"PARK13" "BUS13" "CAB13" "OTRAN13" "ED13" "CHILD13" "HOS13" "DOC13" "PRESCR13" "HINS13"
"HHREP13" "FURN13" "CLOTH13" "TRIPS13" "OTHREC13" "FIPSTATE_DUM1" "FIPSTATE_DUM2"
"FIPSTATE_DUM3" "FIPSTATE_DUM4" "FIPSTATE_DUM5" "FIPSTATE_DUM6" "FIPSTATE_DUM7"
"FIPSTATE_DUM8" "FIPSTATE_DUM9" "FIPSTATE_DUM10" "FIPSTATE_DUM11" "FIPSTATE_DUM12"
"FIPSTATE_DUM13" "FIPSTATE_DUM14" "FIPSTATE_DUM15" "FIPSTATE_DUM16" "FIPSTATE_DUM17"

"FIPSTATE_DUM18" "FIPSTATE_DUM19" "FIPSTATE_DUM20" "FIPSTATE_DUM21" "FIPSTATE_DUM22"
"FIPSTATE_DUM23" "FIPSTATE_DUM24" "FIPSTATE_DUM25" "FIPSTATE_DUM26" "FIPSTATE_DUM27"
"FIPSTATE_DUM28" "FIPSTATE_DUM29" "FIPSTATE_DUM30" "FIPSTATE_DUM31" "FIPSTATE_DUM32"
"FIPSTATE_DUM33" "FIPSTATE_DUM34" "FIPSTATE_DUM35" "FIPSTATE_DUM36" "FIPSTATE_DUM37"
"FIPSTATE_DUM38" "FIPSTATE_DUM39" "FIPSTATE_DUM40" "FIPSTATE_DUM41" "FIPSTATE_DUM42"
"FIPSTATE_DUM43" "FIPSTATE_DUM44" "FIPSTATE_DUM45" "FIPSTATE_DUM46" "FIPSTATE_DUM47"
"FIPSTATE_DUM48" "FIPSTATE_DUM49" "FIPSTATE_DUM50" "FIPSTATE_DUM51" "FIPSTATE_DUM52"
"FIPSTATE_DUM53" "FIPSTATE_DUM54" "FIPSTATE_DUM55" "FIPSTATE_DUM56" "FIPSTATE_DUM57"
"CHGMS_DUM1" "CHGMS_DUM2" "CHGMS_DUM3" "CHGMS_DUM4" "CHGMS_DUM5"
"CHGMS_DUM6" "CHGMS_DUM7" "CHGMS_DUM8" "FCC_DUM1" "FCC_DUM2" "FCC_DUM3"
"FCC_DUM4" "FCC_DUM5" "FCC_DUM6" "FCC_DUM7" "FCC_DUM8" "FCC_DUM9" "FCC_DUM10"
"WF_IND" "OF_IND" "SEX" "AGE" "ES" "BD1_FLAG" "BD2_FLAG" "GEOGMOBIL" "ENRLD"
"GD_ENRLD" "P0P70ACKPT" "P1" "P16" "P20" "P20L" "P20H" "H1HW" "H1AHW" "H1BHW" "H1CHW"
"H2HW" "H3HW" "H4HW" "H5AHWSTROK" "H5BHWHATTA" "H5CHWCORON" "H5DHWBLPRE"
"H5EHWASTHM" "H5FHWLUNG" "H5GHWDIABE" "H5HHWARTH" "H5IHWMEMLO" "H5JHWLEARN"
"H5KHCANCE" "H5LHWNERVE" "H5MHWOTR" "H8HW" "H8ANITES" "COMPED" "KL39" "KL40_1"
"AVGHRS" "TOTOTHR" "WKSILLOTR" "WKSILLSLF" "WKSIVAC" "WKSSTRIKE" "WKSLAYOFF"
"WKSUNEMP" "WKSIOOLF" "AGE_WF" "ES_WF" "BD1_FLAG_WF" "BD2_FLAG_WF" "GEOGMOBIL_WF"
"ENRLD_WF" "GD_ENRLD_WF" "P0P70ACKPT_WF" "P1_WF" "P16_WF" "P20_WF" "P20L_WF"
"P20H_WF" "H1HW_WF" "H1AHW_WF" "H1BHW_WF" "H1CHW_WF" "H2HW_WF" "H3HW_WF"
"H4HW_WF" "H5AHWSTROK_WF" "H5BHWHATTA_WF" "H5CHWCORON_WF" "H5DHWBLPRE_WF"
"H5EHWASTHM_WF" "H5FHWLUNG_WF" "H5GHWDIABE_WF" "H5HHWARTH_WF"
"H5IHWMEMLO_WF" "H5JHWLEARN_WF" "H5KHCANCE_WF" "H5LHWNERVE_WF"
"H5MHWOTR_WF" "H8HW_WF" "H8ANITES_WF" "COMPED_WF" "KL39_WF" "KL40_1_WF"
"AVGHRS_WF" "TOTOTHR_WF" "WKSILLOTR_WF" "WKSILLSLF_WF" "WKSIVAC_WF"
"WKSSTRIKE_WF" "WKSLAYOFF_WF" "WKSUNEMP_WF" "WKSIOOLF_WF"