Online Appendix

Housing Lock: Dutch Evidence on the Impact of Negative Home Equity on Household Mobility*

Asaf Bernstein† Daan Struyven‡

October 1, 2021

*We are extremely grateful to James Poterba, Antoinette Schoar and Bill Wheaton for invaluable advice, guidance, support, and encouragement through this project. We also thank Adrien Auclert, David Autor, Nittai Bergman, Manasi Deshpande, Mathias Dewatripont, François Koulischer, and Christopher Palmer for excellent discussions, comments and feedback. We thank seminar participants at MIT, the Macro Financial Modeling Conference, Berkeley Haas School of Business, Cass Business School, Georgia State University, LSE Geography, Oxford Said, Leeds School of Business, LSE finance, Wharton Business School, the Federal Reserve Board and NYU Stern Finance. Thanks to the Centre for Policy Related Statistics at Statistics Netherlands for providing and documenting the data. We finally thank the MIT Center for Real Estate and the Shultz Fund in the MIT Economics Department for support in purchasing access to the data. All errors are our own.

†University of Colorado at Boulder - Leeds School of Business and NBER; asaf.bernstein@colorado.edu
‡Goldman Sachs economics research; daan.struyven@gmail.com

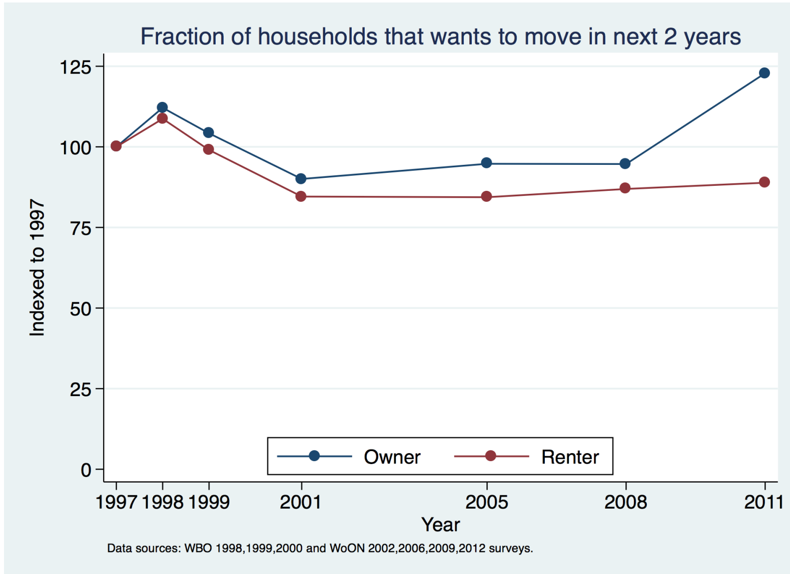

Figure A.1: Mobility intentions of owners and renters over time

Notes: The data are from the WBO 1998, 1999, 2000 and WoON 2002, 2006, 2009 and 2012 surveys. WoON (WoonOnderzoek Nederland) is a repeated cross-sectional nationally representative survey of about 70,000 individuals about their housing situations which was known as WBO (WoningBehoefteOnderzoek) until 2000.

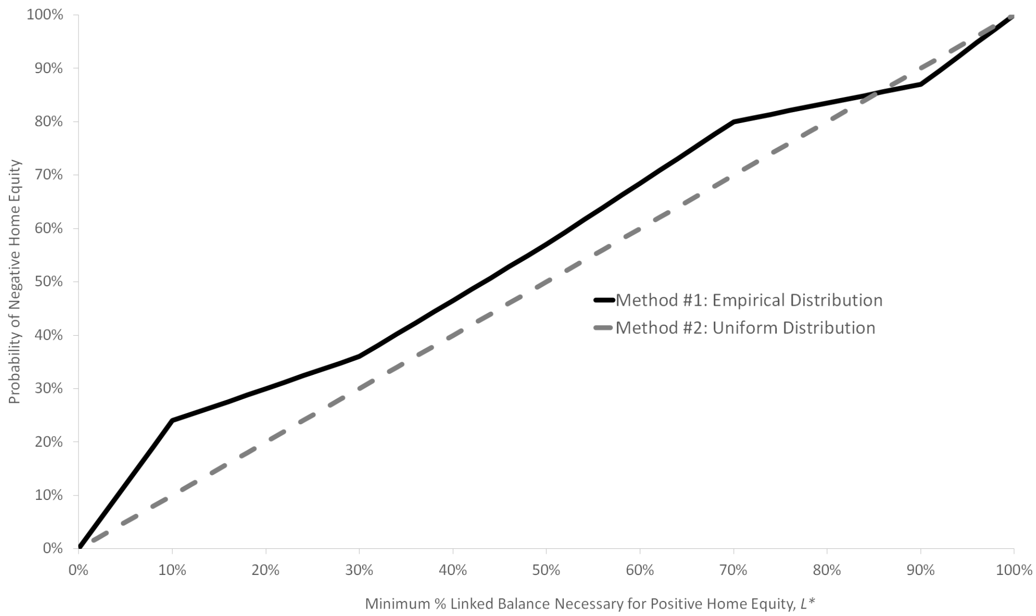

Figure A.2: Probability of Negative Home Equity vs. L^*

Notes: This presents the relationship between the minimum % of the mortgage balance that is linked, therefore amortization in those accounts is not available in our administrative data, necessary for a household to have positive home equity, L^*,and the probability they have negative home equity. $L^*$is computed based on equation 8, and increases as either the reported LTV, or LTV assuming the full balance is linked, rises, since one would have to assume a higher proportion of the balance is linked in order to have positive home equity. Based on these values for L^*, the probability of negative home equity under method #1, the primary method in the paper, is computed following equation 9 and depicted with the solid black line. As L^*increases only households with larger proportions of linked accounts are likely to have positive home equity and so the probability of negative home equity rises, as shown. With the gray dashed line we also depict that same relationship, but using an alternative approach (method #2). In this case we assume a uniform distribution for the proportion of balances that are linked, rather than the empirical distribution.

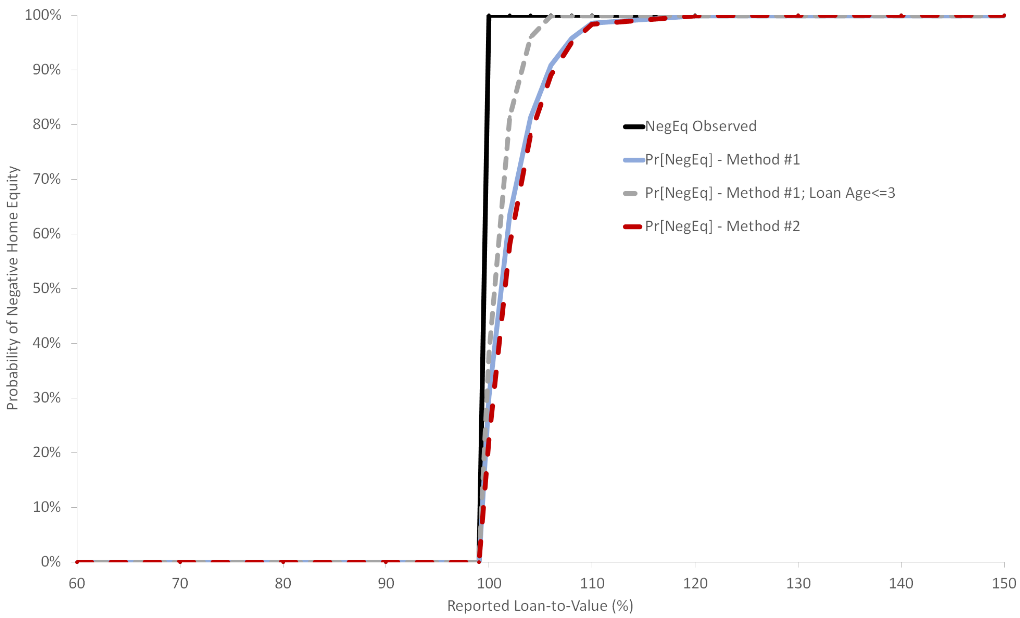

Figure A.3: Probability of Negative Home Equity vs. Reported Loan-to-Value (%)

Notes: This presents the relationship between the loan-to-value as computed based on information reported to our administrative data source, CBS, and the imputed probability of negative home equity under a variety of methods/assumptions. The solid black line depicts a dummy variable equal to one if the reported LTV is greater than 100%. The solid blue line depicts the imputed probability of negative equity using the primary approach (method #1) in our paper following equations 8 and 9. This relationship depends on properties of the loans, including their maturity. The gray dashed line depicts the same relationship, but only for households who moved within the last 3 years. These loans have had less time for unreported amortization to cause a potential difference between reported and actual LTVs, and so are closer to the black solid line. The red dashed line depicts the same thing as the blue solid line, but using an alternative approach (method #2). In this case we assume a uniform distribution for the proportion of balances that are linked, rather than the empirical distribution.

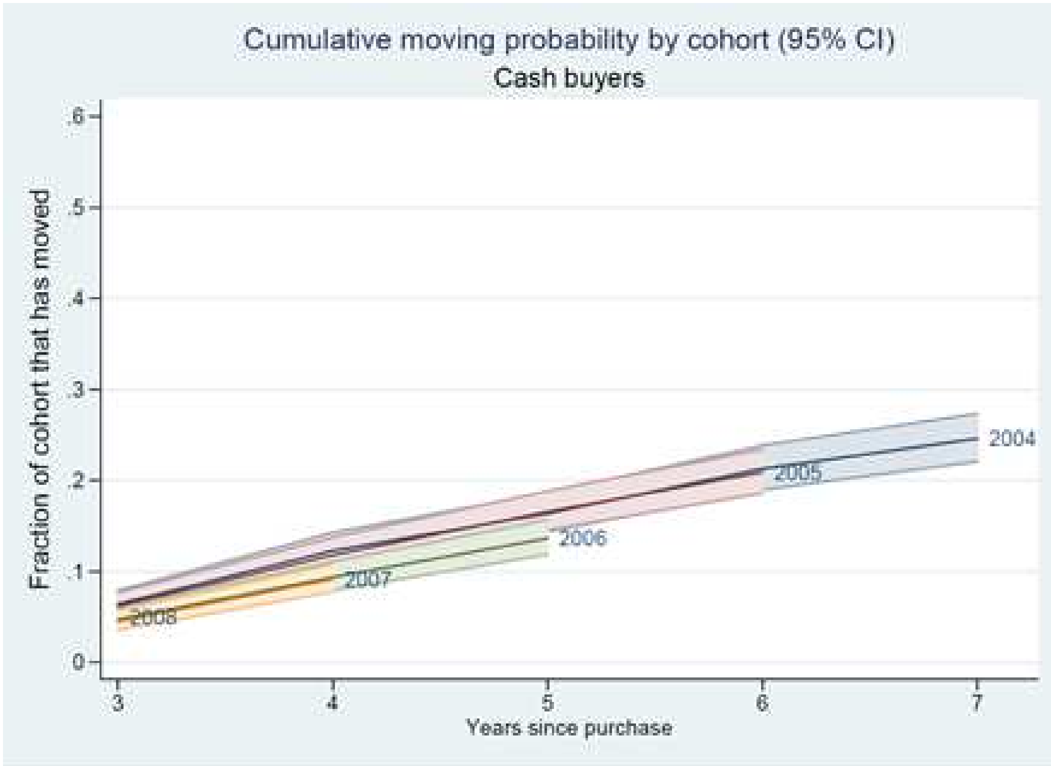

Figure A.4: Mobility of purchase cohorts for cash buyers

Notes: This presents average cumulative moving probabilities and 95% confidence intervals for a placebo test of cash buyers which are differentially exposed to house prices, but who are not exposed to housing lock. The moving data are based on the Transactions Registry and the Address Registry from Statistics Netherlands (CBS). The sample is restricted to transactions of buyers who have no record of mortgage liabilities associated with the home purchase at any date and have balance sheet data over the time period 2007-2012.

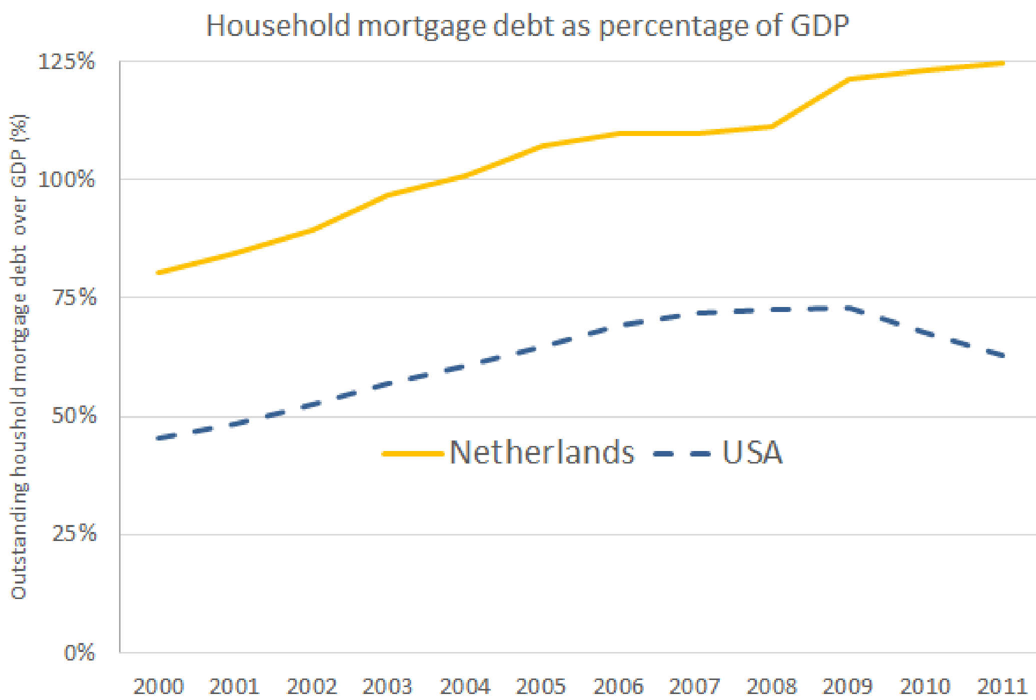

Figure A.5: Household debt to GDP ratio for the US and the Netherlands

Notes: The sources are Statline Statistics Netherlands and Federal Reserve Bank of St. Louis Economic Data.

Figure A.6: Construction of panel of buying heads of existing homes

Notes: The figure provides visual support for the explanation of the construction of the sample of buyers of transacted existing homes and the associated panel of buyer-years. See appendix for more details.

Table A.1: Buyer summary statistics by cohort

Cohort	Age	Household Size	Male	Married	# Transactions
1995	36.00	2.44	0.89	0.51	26261
	(11.49)	(1.19)	(0.32)	(0.50)	
1996	36.11	2.44	0.89	0.50	31916
	(11.41)	(1.20)	(0.32)	(0.50)	
1997	36.10	2.42	0.89	0.48	33271
	(11.41)	(1.19)	(0.32)	(0.50)	
1998	36.09	2.42	0.89	0.47	34331
	(11.31)	(1.20)	(0.32)	(0.50)	
1999	36.27	2.41	0.88	0.46	36540
	(11.42)	(1.20)	(0.32)	(0.50)	
2000	36.38	2.38	0.88	0.44	34431
	(11.55)	(1.19)	(0.33)	(0.50)	
2001	36.68	2.39	0.87	0.43	36482
	(11.69)	(1.19)	(0.33)	(0.50)	
2002	36.67	2.38	0.87	0.41	38001
	(11.65)	(1.19)	(0.34)	(0.49)	
2003	36.86	2.37	0.86	0.41	36634
	(11.73)	(1.20)	(0.35)	(0.49)	
2004	37.20	2.33	0.85	0.40	37217
	(11.98)	(1.19)	(0.36)	(0.49)	
2005	37.53	2.33	0.84	0.39	39387
	(12.03)	(1.19)	(0.37)	(0.49)	
2006	37.76	2.34	0.84	0.38	40933
	(12.08)	(1.20)	(0.37)	(0.49)	
2007	38.02	2.31	0.84	0.38	38757
	(12.42)	(1.18)	(0.37)	(0.48)	
2008	37.59	2.32	0.84	0.37	36472
	(12.32)	(1.18)	(0.37)	(0.48))	
2009	36.25	2.21	0.82	0.33	25933
	(12.22)	(1.14)	(0.38)	(0.47)	
2010	36.78	2.22	0.82	0.33	24538
	(12.31)	(1.15)	(0.38)	(0.47)	
2011	37.65	2.25	0.82	0.35	23233
	(12.74)	(1.16)	(0.38)	(0.48)	

Notes: Age, household size, male and married are measured on December 31st of the year of the move into the property. Male and married are indicator variables for the given characteristic of the buyer. See Section 5 in the text for more details on the CBS data on buyers.

Table A.2: Description of selection of transactions

Criterion	Deleted observations	Remaining observations
Initial transaction data 1995-2011		3,057,528
Missing address	35,242	3,022,286
More than one transaction at address per quarter	32,066	2,990,220
Random 25% selection	2,242,666	747,554
No match to start address spell of person moving in	116,607	630,947
More than one transaction per person per year at address	1,773	629,174
No match to household spell	37	629,137
More than one move per person at a given address in a year	825	628,349
Move after 2011 (we cannot measure subsequent mobility in 2013)	3,034	625,315
Match to exactly one household head	47,500	577,815
More than one purchase of given address by given household head	45	577,770
No match of transaction address to January address of next year	3,433	574,337
Selected transactions matched to head of household		**574,337**

Notes: This table summarizes the number of deleted and remaining observations at each step of the construction of the sample of buyers. See Appendix 7 of the text for more details.

Construction of Sample of Buyers

We construct the 1995-2012 buyers in three steps. First, we randomly select 25% of the transactions of existing owner-occupied homes in 1995-2011. Second, we identify the unique household head among the persons moving into the selected property. Third, we build a panel following these identified buyers over the period 1995-2012. We now detail each of these three steps as well as the number of remaining observations after each step shown in Figure A.2.

Random selection 25% of transactions in 1995-2011. We make use of the universe of 3,057,528 transactions of the existing owner-occupied dwellings file with transaction dates during the period 1995-2011. This file has as identifiers an address and a month of home purchase. We drop 35,242 transactions for which the address variable is missing and keep 3,022,286 transactions**. We** drop 32,066 transactions for which there is more than 1 transaction in a given quarter for a given address and keep 2,990,220 transactions**.** Given memory constraints, we then randomly select 25% of the purchases to obtain 747,554 purchase transactions.

Identifying unique household heads among persons moving in into sampled properties. To identify the unique household head from the persons moving into a selected property, we first consider all the individuals moving in during the same quarter at a given address. We use the universe of individual address spells with coverage January 1995-December 2012, which has as identifiers an address, an encrypted social security number, a spell start date and a spell end date. There are 35,642,414 individual address spells starting after January 1st 1995[23]. We drop 1,198,597 individual address spells with more than 10 persons moving in[24] and obtain 34,443,817 individual address spells. We consider all the address spells on a given address starting in a given quarter and regroup them. The 34,443,817 individual address spells correspond to 21,467,505 household address spells.

We then merge the 747,554 purchase transactions with the 21,467,505 reshaped household address spells using the addresses and quarter of purchase of spell start as keys. 398,826 transactions (=53.35% of the transactions) are matched to a spell that starts in the same quarter as the purchase date. 185,575 (=24.82%) of the transactions are matched to a spell that starts in the quarter after purchase. Finally, 46,546 (=6.23%) of the transactions are matched to a spell that starts two quarters after the purchase. Hence, we match 630,947 of the 747,554 purchase transactions (=84.40%).[25]

To identify the household heads for the selected transactions, we use the head of household identifier

[23] We drop the 15,415,895 left-censored spells starting exactly on January 1st, 1995; the database starts on January 1st 1995.

[24] The main goal is to avoid those who moved into institutional addresses (e.g. senior citizen housing, nursing homes).

[25] For the remaining 116,607 non-matched transactions, **we** observe the variable "Is the buyer a current renter" (which has always been measured since 1998 and never before) for 85,732 transactions. 27,276 of those 85,732 non-matched transactions are bought by current renters (17,028 are sales by public housing corporations).

dummy on December 31st of the transaction year from household structure spells that we will match to transactions. The head of household dummy is created by Statistics Netherlands with a time-consistent and intuitive rule. If there is a couple in the household, then the male member of the couple is the head of household. If the couple is of the same gender, then it is the oldest person. The head of single-parent household heads is the parent. In an "other household", the head is the oldest male, 15 years or older- and if this is missing- the oldest woman, 15 years or older. In multiple-generation households (e.g. non-married pair with daughter and mother), then the partner in the couple rule dominates the parent rule in a single-parent family and in the case of two (via child-parent related) pairs, the head is chosen as the youngest pair.

To select the head of household from the persons who moved in an address, we list all the individuals moving into the sampled transactions, and we consider 1,441,087 person moving-in years corresponding to the 630,947 transactions. The 1,441,087 person moving-in years correspond to 1,364,733 distinct persons. Focusing on persons who have only 1 transaction per year per address in our 630,947 transactions, we drop 4,445 person-years and get 1,436,642 moving in person years, which corresponds to 1,362,594 distinct persons and 629,174 transactions. To identify the head in the year of the move, we then build a annual panel of household structures of persons moving in. We therefore merge the universe of 138,238,794 household spells with the list of 1,362,594 transacting distinct persons using the SSN. Household spells have as a unit of observation a SSN, a family-structure-spell start date, a family-structure-spell end date and a household number. 21 persons cannot be matched and we find 13,244,303 individual household structure spells for the 1,362,573 matched distinct persons. Because of insufficient disk space constraints, we drop the 1,576 household spells (0.01% of 13,244,303 spells) with more than 150 household spells to keep 13,242,727 spells. We then match the moving in person-years and person-years from the household structure panel. From the 1,436,642 moving in person years (629,174 transactions), we can match 1,436,459 person years (629,137 transactions) to their household structure in December of the year of the transaction. We then restrict ourselves to the 1,434,705 moving in person years (628,349 transactions) where there is only 1 selected move in that year for that person at that address. We then drop transactions for which the move starts in 2012, as we cannot observe subsequent mobility out of purchase dwellings in 2013 and later as of yet. We thus drop 6,750 moving in person years for which the moving in date occurs in 2012 (0.47%), and we obtain 1,427,955 moving in person-years (625,315 transactions). We then match transactions and heads.

From the 625,315 transactions, we can match 577,815 transactions (=92.40%) to exactly 1 household head (as defined by SN) using the panel of household structures of persons moving in. However, 5.06% of the transactions have no household head and 2.20% of the transactions are associated to starting address spells for 2 household heads. To keep things simple and non arbitrary, we keep the 577,815 transactions associated to starting address spells with exactly one household head (which corresponds to 1,332,388 moving in person years). The 577,815 transactions- that we can match to the start of the address spell of a unique household head correspond to 552,168 distinct persons. 527,407 (95.52%) persons occur once, 23,912 persons twice (4.33%), 812 persons three times (0.15%) and 37 persons four times (0.01%). We then use the universe of 2012 time-unvarying personal characteristics file GBAPERSOONTAB and merge it with the list of 577,815 selected and matched buying heads of households using the SSN as key.

Building a 1995-2012 panel for selected buyers To know the address before and after the purchase, we build a panel of December addresses for the 552,168 distinct persons retaining the 2,175,981 address spells of the 552,168 distinct persons. We reshape the 2,175,981 address spells into 552,168 lines where we put the 1 to 40 addresses of a given person on 1 line. We then reshape the file to create 18 December addresses for the 552,168 distinct persons which corresponds to 9,939,024 person-years (=18*552,168). We then merge the 9,939,024 person-years and the list of 552,168 distinct persons using as key, the SSN and the year (where the year is the year in which the address spell associated to the transaction began). Finally, we implement two minor transaction sample restrictions using the panel of addresses. First, before defining mobility, we drop 378 person years- which corresponds to 45 transactions- if the same person buys the same address more than once to keep 9,938,646 person years (552,147 persons) and 577,770 transactions. Second, for 3,433 out of the 577,770 transactions, the January address after the move in is already different from the address where the buyer moved in. We focus on the remaining 99.44% or 574,337 transactions.

Dutch Mortgage Characteristics and Costs of Housing Lock

Dutch Mortgage debt stock and characteristics The current Dutch residential mortgage-to-GDP ratio of approximately 120% is the highest in the world, which is approximately 45 percentage points higher than in the US, as shown in Appendix Figure A.5. This high mortgage-to-income ratio reflects (1) high home-price-to-income ratios, (2) a moderate homeowmership rate, and (3) high LTV ratios among homeowners.

First, the median Dutch household housing cost burden in 2014 is 23.9% of disposable income. This average housing cost among owners and renters corresponds to percentile 90 among OECD countries (vs. 18.9% OECD average and 19.5% in the US). Second, the homeownership rate in the Netherlands is 60%. Third, average Dutch LTV ratios are high, reflecting both high LTV ratios at origination and limited amortization. LTV ratios at origination around 100 or even slightly above 100% are not unusual in the Netherlands. In the latter case, the loan proceeds can finance the entire purchase price of the house, transaction costs such as the 6% stamp duty (reduced to 2% in July 2011) or home improvements.

The vast majority of mortgages for the 1995-2011 purchase cohorts that we study are non-amortizing. Interest-only loans are frequently combined with associated, pledged accounts where capital is built up in the form of savings deposits, life insurance or investment funds. Mortgage contracts often combine multiple loans with different repayment types, for instance a plain vanilla interest-only loan, combined with a second interest-only loan with an associated savings deposit account. Contracts with associated tax-exempt accounts allow borrowers to build up capital while maximizing the unlimited deduction of interest payments on the constant loan balance.[26] As owner-occupied homes are considered a source of income, an imputed rental income of 0.6% of the value of the house is included in taxable income. Relative to the US, both relatively high marginal tax rates on personal income, that rise from 36 to 42% at €19,646 of taxable income and to 52% at €56,532 of taxable income[27] and the absence of the itemizing precondition for claiming the deduction, increase the economic importance of the deduction. The typical mortgage features a maturity of 30 years and an interest rate that is fixed for 10 years and then periodically reset.

The cost of housing lock Declines in residential mobility can in principle be associated with several costs, including, for instance, lower job switching and job quality, declines in the quality of housing matches, lower ability to smooth income risk, and reduced performance at existing employment. Several authors have explored such factors in the U.S. setting including Brown and Matsa (2016), Demyanyk et al. (2017),Bernstein, McQuade and Townsend (2017) ,and Gopalan et al. (2017).The magnitude and relative importance of these costs likely also depends on policy factors. In terms of the baseline residential mobility rate, the mobility of Dutch homeowners is comparable to US levels (Emrath (2009)) with estimated times until half of the buyers move from their homes of around 12 to 13 years. First, in terms of the labor market effects, the Netherlands is of course a smaller country than the US with more synchronized local business cycles. But within the European Union, the Netherlands is classified as a high-geographical and high-job-mobility country, together with the UK, the Scandinavian and Baltic states (Vandenbrande, Coppin and Van der Hallen (2006)). The average Dutch job duration is approximately 6 years compared to 8 years in the EU with shorter durations only for Denmark, the UK, Latvia and Lithuania. Second, the comparable baseline homeowner mobility rate suggest that the cost from foregone moves with non-labor market motives (e.g. proximity to family, change in house size) is likely of a similar order of magnitude. Third, the cost of the inability to smooth income risk by adjusting housing costs probably also depends on the availability of social insurance–perceived as relatively generous in the Netherlands–and the options to default, which are limited in the Netherlands. Overall, the Dutch institutions allow isolating the housing lock, likely imply a more negative effect of home equity on mobility than in the US, and have led to a very high mortgage-debt-to GDP ratio.

[26] As of January 2013, new mortgages have to fully amortize to benefit from interest tax deduction, which has decimated the market for non-amortizing loans (Struyven (2015)).

[27] The maximum rate for interest deduction is reduced gradually since 2014 from 52% to 38% by 50 basis points a year.

Toy Model on the interaction of housing lock and moving distance of buyers

The Stein (1995) model predicts that households move when (i) the life-time utility benefit of moving exceeds the life-time utility cost of moving, and (ii) the total net liquid assets are larger than the required downpayment and the pecuniary upfront moving cost (i.e. when the liquidity constraint does not bind). As both the life-time benefits (ex. income gains from job search) and costs (ex. shipping costs, distance to family, job search costs) vary with distance, it is not obvious theoretically how the effect of negative home equity on mobility should vary with distance.

In terms of the model, let us define B as the discounted present value of the gross utility benefits of moving and C as the discounted present value of the gross utility costs of moving. In turn, the discounted costs C are the sum of (1) pecuniary moving costs MC faced today (e.g. trucks, furniture, time away from job), and (2) other costs OC. Additionally, the household balance sheet consists of home equity denoted by HE , and liquid financial assets FA. Let us also assume that a moving household has to put down D but can obtain a supplementary personal loan L if it moves (by borrowing for instance against higher future expected wages). Finally, the household also keeps a minimum precautionary buffer P of cash set aside (i.e. the minimum on their cash account), that may increase in a new, riskier environment.

We can posit that a household moves if a new home opportunity arises for which (i) the life-time benefit of moving exceeds the life-time cost of moving:

$$B > C \tag{10}$$

and (ii) the total liquidity after the move exceeds the minimum precautionary buffer:

$$(HE + FA + L) - (D + MC) > P \tag{11}$$

Let us assume that the household has a short-distance moving opportunity s and long–distance opportunity l. We also assume that the discounted moving benefit $B(d)$, the pecuniary moving cost $MC(d)$, the discounted total moving cost $C(d)$, the collateral value of moving $L(d)$, and the precautionary buffer $P(d)$ can all vary with the moving distance d, while home equity HE and financial assets FA–predetermined before the moving opportunity arises–do not depend on distance. Specifically, it is plausible to assume that $B(d)$ rises in distance as more job more opportunities and person-specific amenities–such as weather–become available. On the cost side, DaVanzo (1981) finds that individuals build location-specific capital, and Mulder and Wagner (2012) and Mulder and Malmberg (2011) show that moving tends to occur less frequently if a person has family nearby and moving probability and distance are lower if the person has lived in a location for a long period of time.

How does the impact of negative home equity on mobility differ for the short and long moves? In this simple static setting negative home equity only affects equation 11. Negative home equity has a larger effect on the long move than on the short move if equation 11 is more often violated for long moves l than for short moves s. The two threshold conditions for this to be true are (i) $(HE(s) + FA(s) + L(s)) - (D(s) + MC(s)) > P(s)$ and (ii) $(HE(l) + FA(l) + L(l)) - (D(l) + MC(l)) = P(l)$. Combing these 2 conditions and using that home equity, financial assets, and the downpayment are equal for short and long moves gives the following condition:

$$L(l) - L(s) < [MC(l) - MC(s)] + [P(l) - P(s)] \tag{12}$$

In words, negative home equity has a more negative effect on long distance moves than on short distance moves if the extra collateral value of a long move is smaller than the sum of the extra moving costs and the extra precautionary cash buffer. This would happen if the labor benefits of moving are sufficiently high and can be easily borrowed against for farther moves. In the absence of the ability to borrow against future expected earnings this static model would predict that farther moves would tend to be more sensitive than shorter moves to negative home equity.

Linked Accounts and Computation of Negative Home Equity Probabilities

As we note in section 5, some borrowers have "linked" mortgage-savings accounts, which complicates the computation of negative home equity. These borrowers have partially amortizing mortgages, but instead of reducing the mortgage balance in the official tax records, they instead build-up money in a sinking fund that can only be used for paying off the mortgage. These accounts are tax exempt and are therefore not observed in our tax records for either the mortgage balance outstanding nor listed as other forms of financial assets. This means that some households that appear to have negative home equity from our tax records, actually do not have negative equity. If they wanted to move, the mortgage balance is actually lower than it appears, by the amount of the amortization built-up in the linked accounts.

Not accounting for these linked accounts could be potentially problematic within our 2SLS framework. For the sake of simplicity, imagine that hypothetically there is a valid instrumental variable we can construct, where for an "as if random" subset of borrowers the mortgage balance is increased. Negative home equity is our treatment of interest, so some, but not all of the borrowers who receive this increase in mortgage balance will be treated, but in general a higher mortgage balance will increase the probability that a given borrower has negative home equity. For ease of exposition, and building off the language of intent-totreat (ITT) analysis, lets call borrowers who are not assigned to receive an increase in mortgage balance, "unassigned borrowers", and those that are "assigned borrowers." Also assume that no unassigned borrowers have negative home equity, while 25% of assigned borrowers have negative home equity. In that case the true 1st stage for the effect of treatment on negative home equity would be 25%, $\gamma = 0.25$. Let the true effect of negative home equity on moving rates be $\beta = -4\%$. In that case the true reduced form coefficient δ would be -1%, since $\delta = \gamma\beta = 0.25 \times -0.04$. Now assume that we can observe true assignment and outcomes, but not actual treatment, and that hypothetically it is the case that 50% of borrowers with observed negative home equity in reality have positive home equity, if linked accounts could be observed. In that case, since we can only use the observed negative home equity, we would estimate that $E[\hat{\gamma}] = 0.5$, since it would appear to the econometrician that half of assigned borrowers have negative home equity, when the reality is that only 25% do. The econometrician observes true assignment and outcomes and the reduced form estimate would be unbiased, $E[\hat{\delta}] = \delta = -0.01$. Putting these together though the 2SLS estimate would be biased since $E[\hat{\beta}] = \frac{E[\hat{\delta}]}{E[\hat{\gamma}]} = \frac{-0.01}{-0.5} = -2\%$. In other words, asymmetric bias in the measurement of the endogenous variable of interest (in this case negative home equity) causes bias in the estimates obtained using 2SLS.

One solution to this is to correctly note that the true observation of the econometrician isn't a binary variable for whether a household has negative home equity or not, but rather an indication of the probability a given household has negative home equity. In other words, in the hypothetical example above, if a borrower was observed to have negative home equity, the econometrician could deduce that they actually have a 50% chance of truly having negative home equity. Replacing the "1s" in the regression with "0.5" would more accurately reflect the true data and in this case undo the bias. In particular, the first stage estimate would now become $E[\hat{\gamma}] = 0.25 = \gamma$ and so the estimate from 2SLS would map back into the true value of the relationship, $E[\hat{\beta}] = \frac{-0.01}{-0.25} = -4\% = \beta$. This same logic can be extended more generally to instances where there are observations for the LTV and a range of amortizations unobserved at the individual level, but understood in aggregate.

In particular, we use exactly this approach when running our analysis, so that instead of using binary variables for negative home equity, we compute the probability of having negative home equity, conditional on the observed LTV of the borrower and information on the aggregate distributional properties on the % of mortgage balances that are partially unobservably amortizing and by what amount. Mastrogiacomo and Van der Molen (2015) provide aggregate statistics on not only the % of mortgages that have unobserved amortizing mortgages, but also what % of mortgages have unobserved partial amortizing for varying proportions of the total mortgage. So in practice what this means is that if we observe a mortgage that appears to have an LTV of 115% and originated their mortgage 50 months ago, we know there is some probability they have unobserved amortization, which means that LTV would be lower, and then conditional on having unobserved amortization there is some chance that 5%, 10%, 25%, 50%, etc of the total balance is made of this unobserved amortizing loan. The higher the proportion of the loan which is unobservably amortizing and the longer since the beginning of the mortgage, the lower will be the LTV of the borrower. So in the case of the borrower with an observed LTV of 115% and a mortgage aged 50 months we would compute the largest % of the mortgage that could be unobservably amortizing over 50 months such that the LTV would be less than 100%. We would then use the aggregate statistics from Mastrogiacomo and Van der Molen (2015) to measure the % of mortgages likely to be unobservably amortizing by that amount and one minus that would be the probability that this borrower has negative home equity. For mortgages where the borrower would have negative home equity even if 100% of the mortgage balance were unobservably amortizing, which is likely for newer mortgages and those with high LTVs, then the probability of negative home equity is treated as 100%. Throughout our analysis we are careful to show that our estimates are not that sensitive to the exact unobservable amortization assumptions used or relying on just borrowers were the endogenous variable is invariant to the treatment of unobservable amortization because they are sufficiently new and/or have high enough LTVs.