DNB Working Paper
No 824/December 2024

Residual Mortgage Debt, Insurance, and Defaults in the Netherlands

Madi Mangan, Mauro Mastrogiacomo and Hans Bloemen

DeNederlandscheBank
EUROSYSTEEM

Residual Mortgage Debt, Insurance, and Defaults in the Netherlands

Madi Mangan, Mauro Mastrogiacomo and Hans Bloemen *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 824

December 2024

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Residual Mortgage Debt, Insurance, and Defaults in the Netherlands

Madi Mangan[a,b,d], Mauro Mastrogiacomo[a,c,d] & Hans Bloemen[a,d]

a: Vrije Universiteit of Amsterdam (VUA), b: Ca' Foscari University of Venice
c: De Nederlandsche Bank (DNB) d: Network for Studies on Pension Ageing and Retirement (NETSPAR)

March 2024

Abstract

Mortgage defaults are commonly linked to affordability and borrowers' income; less often, to a decrease in home value. However, some studies talk about "strategic defaults", a form of moral hazard whereby people who can afford their underwater mortgage choose not to pay. In this way, they clear their excess debt, as single recourse systems act as insurance. Our focus is on a type of mortgage insurance, available for houses with values below a certain threshold, that varies over time. We examine how this mortgage insurance affects decisions to default. We combine a quasi-natural experiment with the estimation of a structural model, more precisely an optimal stopping model. Our findings reveal that the (utility from) future value of home equity negatively influences the likelihood of default. We show that the discontinuity around the qualification threshold is linked to borrowers' income, due to loan-to-income caps. The model indicates that while the insurance does not cause defaults in general, it does lead to more defaults for borrowers who separate from their partners, possibly indicating moral hazard.

JEL codes: G11, G21, G52
Keywords: Residual Mortgage Insurance, Non-performance, Structural model, quasi-natural experiment.

Corresponding author Mauro Mastrogiacomo; address: De Nederlandsche Bank (DNB), P.O. Box 98, 1000 AB Amsterdam, The Netherlands, email: M.Mastrogiacomo@dnb.nl.
The authors acknowledge the financial support of Netspar. Any opinions expressed are those of the authors and not necessarily those of Netspar. We would like to thank Giacomo Pasini, Agar Brugiavini and the other participants of at the seminars organized by Ca' Foscari University of Venice and DNB.

1 Introduction

Fluctuations in house prices can result in negative home equity and an increase in mortgage defaults. This was the case both in the Great Depression (1929 – 1939) and the Great Recession (2007 – 2009) in many economically developed countries. Yet, this link is not immediately obvious. In stress-test models, the probability of default typically depends on affordability (income and mortgage premiums) and not on house values (that determine the loss given default). To rationalize an increase in defaults in the presence of lower house prices, some authors refer to strategic defaults (Campbell et al., 2009), the defaulting of those who could afford the payment of mortgage premiums but prefer to clear their debt. Internationally, the drop in home equity is not always associated with strategic defaults (Bajari et al., 2008). For instance, in the US, defaults increased substantially during the great recession (Mayer et al., 2014), and this increase was partly strategic (Guiso et al., 2013; Gerardi et al., 2018).[1] But in several EU countries, such as Denmark and the Netherlands, there was hardly any increase in the default rate despite a sharp drop in prices.

Institutional and behavioural differences can explain some of these divergences. A single recourse system in the US, but absent in the EU, can act as insurance for underwater mortgages and possibly induce moral hazard. The onset of negative home equity with such a recourse system is regarded as key in the default process (Mian and Sufi, 2013). So, the case of the Netherlands is particularly interesting, as a residual mortgage debt insurance is in place, but with full recourse. This raises the question whether also this insurance induces defaults and could be linked to moral hazard. This (government-backed) insurance (Nationale Hypotheek Garantie or NHG) is only accessible for those purchasing a home with a value below a given threshold, exogenously set (in the sense that is not affected by the borrower), and changing over time. This offers a unique quasi-natural setting to identify the relevance of debt insurance in mortgage defaults.

When one evaluates insurance, moral hazard and/or adverse selection (Marshall, 1976) could be difficult to separately observe/identify (Mayer et al., 2014). But the popularity of the NHG during the great recession excludes adverse selection, so if an effect of insurance on defaults emerges, this could be ascribed to moral hazard.

We notice a negative gradient between house value and mortgage defaults only for NHG participants in the Netherlands. As qualification criteria for debt reimbursement (divorce of the borrowers for instance) occur more often among NHG participants (Kim et al., 2022), in the present study we investigate whether we also observe more defaults among those insured. Put differently, has the insurance, originally set up to protect the mortgage market from idiosyncratic shocks, also induced strategic defaults? To preview the results of the reduced form analysis, we find that mortgage insurance in the Netherlands did not induce strategic defaults. This is consistent with the earlier results for the EU countries (Serrano, 2005) and the Netherlands in particular. The negative

[1] Industry-based analyses (Riley, 2013) of credit bureau data in the United States conducted between 2009 and 2011 suggest that between 12% and 19% of all mortgages, defaults have been strategic in recent years.

gradient with home equity can be explained by confounding factors, and we do not observe a drop in defaults at the insurance threshold.

We complement the quasi-experimental setting with a more structural approach. First, we set up a dynamic programming model (Rust, 1989), which we also tentatively estimate. However, our data does not support the backward recursion solution required by these models. We lack information about the entire mortgage contract duration and the account status in each period leading up to maturity. Our observations cover the initial 10 years after origination, which is a remarkable advancement in data collection for the Netherlands, but not yet enough for contracts that typically last about 30 years. Consequently, we need to estimate a forward-looking model, relying on assumptions about the missing information for the period eliciting until maturity. A useful tool in this context is the option value model, which explains optimal stopping. A recognized application of this model is in the realm of retirement decisions (Lumsdaine, Stock, and Wise, 1992).

According to the Option Value Theory (Foster and Van Order, 1984 and 1985), one might choose to default if the lifetime utility of this decision is higher relative to the option to stay performing. At every payment, the borrower decides whether to default or not. The decision to default does not directly depend on current borrower's income, but rather depends on the borrower optimising utility over time. Default is therefore the result of rational choices of borrowers, with negative equity. To this theory, (Foster and Van Order, 1984, 1985), all underwater mortgages will eventually default. However, evidence suggests that this is not the case and that default can also be due to other factors.

Teye et al. (2017) highlight that the probability of default depends on the characteristics of the mortgage at origination and the composition of the household at the time of default. A high level of indebtedness associated to low affordability, or low household income, can as well trigger defaults (Li and Oswald, 2017). Income and house price changes explain non-performance (Haan and Mastrogiacomo, 2020), along with excessive indebtedness (Li and Oswald, 2017), too high LTV and DSTI caps (Geanakoplos (2010), Wong et al. (2011) and Gerlach-Kristen and Lyons (2018)) and non-recourse (Guiso et al. (2013), Gerardi et al. (2018), Riley (2013), Mayer et al. (2014)). Studying how consumer utility develops in the presence of mortgage insurance is particularly interesting. Typically defaulting on a mortgage should be expensive to the borrower, in the presence of a full recourse mortgage (Hatchondo et al., 2015), as non-performance lowers home equity. However, with residual debt insurance such as the NHG, default is less costly to the borrower (Haan and Mastrogiacomo, 2020).

In our structural model of default choices, we account for this by computing these costs, and those of the alternative option, that is renting a new property. The model describes the utility of revealed preferences, and our results show that consumers derive positive marginal utility of consumption, in line with theory (Campbell and Cocco, 2015). In a discrete choice model this implies that the choice to stay performing delivers a higher utility than defaulting. Our results also show that this is more so in the presence of NHG. This result is also confirmed by the discontinuity analysis, with lower default for those insured and nearing the threshold.

Further, we found that NHG has induced more divorces (Kim et al, 2022), but less defaults. This is not a surprising result, as the insurance becomes active when someone divorces (or gets unemployed or disabled) but does not condition the reimbursement of residual debt on defaulting. Yet, when we restrict our analysis to those becoming singles, our results indicate that while the insurance does not cause defaults in general, it does lead to more defaults for borrowers who separate from their partners, possibly indicating moral hazard. So, the insurance protects the financial system well, but not necessarily taxpayers, as these are in the end potentially liable should the fund that governs the NHG run in financial troubles.

We explain how we got to these conclusions in the remainder of this paper. The rest of the paper is organized as follows. Section 2 describes the main institutions operating in the Dutch housing and mortgage market. Section 3 discusses our data and presents some relevant descriptive statistics. Section 4 introduces our empirical model. Section 5 contains our main results and robustness checks.

2 Institutional context

2.1 The mortgage market in the Netherlands

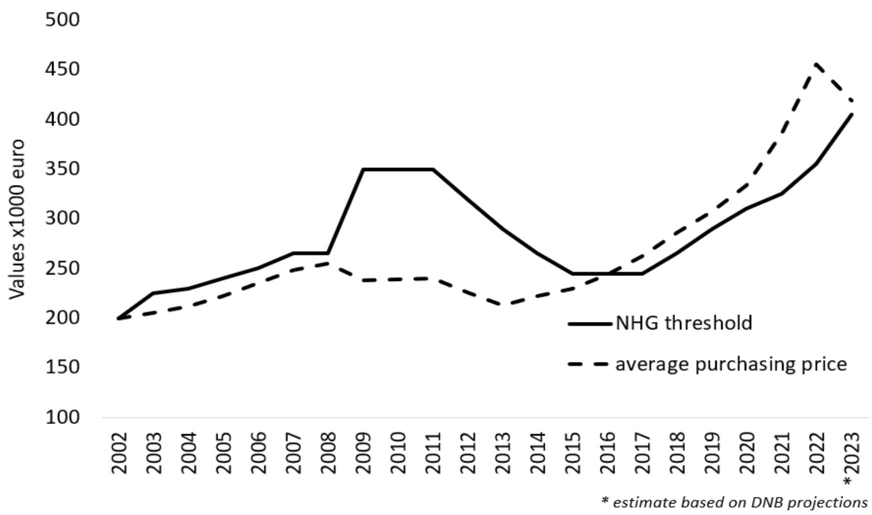
The Dutch mortgage market is highly concentrated. The three largest banks provide the largest share of all loans (about 3/4), though non-bank lending, mostly by insurers and pension funds, has been increasing in size in the past decade. As these institutions securitize, they are requested to report information about their mortgage portfolio, loan by loan, to the relevant supervisors. This is the data source of the present study. The mortgage market is characterized by low non-performance rates, although it has among the highest loan-to-value (LTV) and loan-to-income (LTI) ratios at origination in Europe. LTV ratios are currently capped by law at 100%, after a lengthy process of reduction that took almost a decade. LTI-equivalent DSTI caps are yearly set, but these are complied with or explain rules. House price is highly unstable in the Netherlands. Experiencing its lowest price growth in 2013, house prices have continued to rise in the Netherlands since then except for the recent decline in the last quarter of 2022 and first quarter of 2023.

Another distinctive feature is the high share of interest-only (IO) loans, about half of the portfolio now, which has attracted the attention of (international) supervisors because of financial stability risks (ESRB, 2022). As of 2013, this type of loan is discouraged in the tax system, but this fiscal disincentive had lost effectiveness due to the low-interest rates that had been offered till mid-2022. The high share is though mostly a legacy of the past, and the sharpening of risk weights for these products should further discourage their production. A full recourse system exists in the Netherlands, and it implies unlimited liabilities to home buyers with mortgage loans.

2.2 NHG

The residual mortgage debt insurance (Nationale Hypotheek Garantie or NHG) is a distinctive macro-prudential tool backed by the Dutch government. Qualification for the insurance depends on a threshold for the purchasing price set yearly (below €435,000 in 2024). Homes bought below this limit are eligible for insurance, while pricier homes cannot be insured. The threshold is set by the fund overseeing NHG in coordination with the Dutch government. It is important to note that this limit does not follow a predictable pattern. Before the Great Recession, it was slightly above average purchase prices, significantly surpassed average prices during that crisis, and eventually settled below (see Figure 1). Notice that during the Great Recession about 40% of all Dutch mortgages were underwater. This exogenous and unpredictable variation will be beneficial for the discontinuity design that we present later.

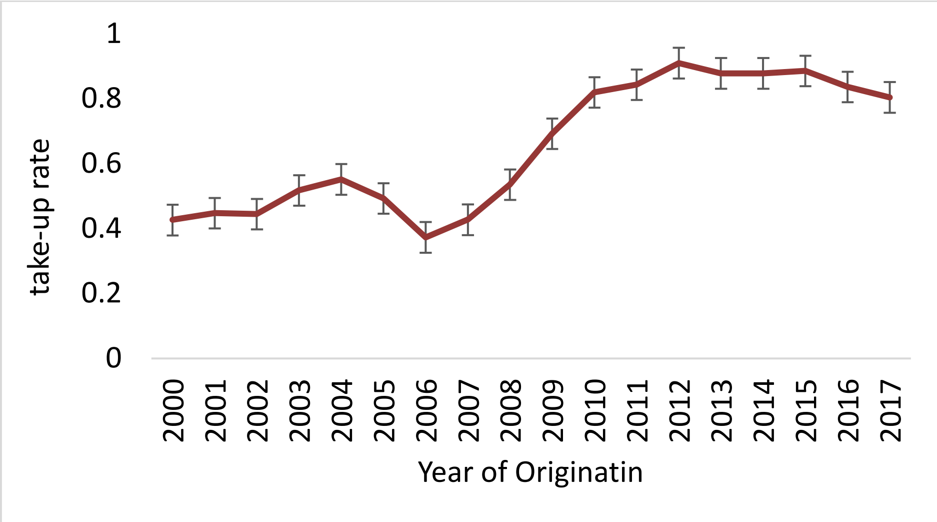
Figure 1: NHG threshold and average purchase price, 2002-2023



Explanatory notes: prices in 2023 are estimated on the base of DNB forecasts. In 2023 for the first time, there will be a higher threshold (about 430k) for more energy-saving houses. Threshold values are reported on December 31st of each year (though these might have changed at different months in that year).

NHG was immensely popular among first-time home buyers in the years around the Great Financial Crisis (see Figure 2, not unrelated to NHG threshold being relatively high to the average purchasing price, Figure 1), its popularity has decreased with the lowering of the interest rates and the increase in house prices witnessed till 2022. Its popularity depended on two reasons: it serves as insurance for borrowers, and it allows them to borrow at the lowest available interest rate, even if one borrows at high LTV ratio's. At the same time, NHG is a macroprudential tool that helps prevent possible financial instabilities that would stem from macroeconomic shocks, possibly causing more defaults. This is why also mortgage providers are happy to offer it.

Figure 2: Uptake rate of NHG conditional on qualification by year of origination



Explanatory notes: Uptake rates are computed for mortgages with a value below the NHG threshold and originated in the corresponding year. With LTV of 90% or less. We focus on new mortgage production by also young borrowers.

Borrowers currently pay 0.6% of the mortgage principal as a one-off entry fee at the inception of the mortgage. They may choose to not buy the insurance if they do not see the risk of their mortgage ever going underwater or their household experiencing a shock which may induce defaulting. Also, some (myopic) borrowers may renounce buying NHG (and even the mortgage interest deduction) to purchase a fully interest-only loan as the two cannot coexist. In the years around the great recession however the NHG was an offer one could not refuse (Haan and Mastrogiacomo, 2020).

It is worth mentioning, reimbursements do not require borrowers to default, but if they do, their residual debt is taken on by NHG. The fund will do so whenever there is a death of one of the partners, divorce, disability of one or both partners, or unemployment, whether this results in default or not. So, differently from the US, one cannot clear debt only because the mortgage is underwater.

Part of our estimation is based on the institutional discontinuity at the threshold, whereby those who do not qualify for NHG cannot insure against default risk. We claim this discontinuity to be exogenous, as the NHG threshold is set yearly and does not depend on individual choices. In the future the threshold will follow average house prices (NIBUD, 2022), but so far this has not necessarily been the case. For instance, in 2023 house prices fell but the NHG threshold increased, and it was even allowed to diverge for more energy-saving houses for the first time.

3 Data and descriptive evidence

We link Dutch Central Bank (De Nederlandsche Bank – DNB) Loan-Level Data (LLD) to different administrative register data and tax records. The Loan Level Data (LLD) is a data set on household mortgages collected quarterly and contains 80 per cent of all residential mortgages in the Netherlands present within securitizing institutions between the last quarter of 2012 (2012Q4) and the final quarter of 2018 (2018Q4). However, because the dataset is also retrospective, some information on the loans and mortgage such as income of the borrowers, LTI, LTV age of the older borrower, employment of status, and other conditions of the loans, can be recovered from the date of origination. It collects information on three (3) million borrowers and their six million loans. The LLD has many unique features which made it suitable for our purpose, as it provides key details which would not otherwise be available, such as participation in the NHG and default status.

We combine the last quarter of the LLD in each year from 2012 to 2018 and merge it to tax and registry data from Netherlands Statistics (CBS)[2], thereby adding data on household/individual income and wealth, the household composition, cohabitation status, and other socioeconomic characteristics of the household such as employment status. This information will help in the identification of the effect of household shocks and other informal debt-mitigation strategies available to the household.

The original dataset is an unbalanced panel: households with defaulted mortgage eventually exit the dataset and households with new mortgages are admitted. When borrowers migrate across banks, we can only track those that are merged to tax records (fortunately the majority).

Despite the unique and valuable features in the data, there is a drawback for our analysis. Ideally, we would want to observe the complete lifecycle of a mortgage, from its start to maturity, and track its (default) status in each period. However, we currently have data for only 25 quarters (2012Q4-2018Q4). To extend this period, we include retrospective information from 2012 down to 2007, considering that in the Netherlands, loan renegotiations rarely occur before 5 years from origination. This is due to the fixation period of interest rates, which is seldom below 5 years, and revolving rates are extremely scarce. Therefore, we can examine the cohort of loans originated in 2007, which did not default before 2012 and still existed in 2018. If a loan or a mortgage is defaulted between 2012 and 2018, the LLD seize to collect information on the succeeding years after default. (Loans that are defaulted exit the LLD a year after default). This gives us information on the development of these loans over 11 years, with most defaults occurring between 2012 and 2018. While we could push back some defaults in time based on default dates, this is only possible if banks retained these defaulted loans on their books until 2012. Defaulting in the initial 5 years after origination is highly unlikely, as borrower solvency is assessed during origination. Consequently, we utilize a panel spanning 11 years, with attrition occurring only due to defaults.

The administrative data originate from the tax office, providing details at both the household and individual levels, including income and wealth. Mortgage loan information is preloaded, with

[2] Household income data, Household wealth, and partnership registers all from the CBS

the mortgage servicer delivering data at the loan level to the tax office. When the servicer is a securitizing institution, like a bank, this information aligns with part of the details in the LLD, enabling a match. Netherlands Statistics managed this matching process, and the resulting information can be further merged with registry data, allowing us to gather details on household composition and cohabitation status. Despite the intricacies of this unique merging process, approximately 80% of the LLD data can be successfully merged with the administrative files for the years 2016 and 2017, for other years the percentage successfully merged decline as we go farther from the year 2016.

3.1 Loan characteristics

The highest level of granularity here is at the loan level. On average, Dutch borrowers have two loans in their mortgage. All characteristics of the loan are compiled from the LLD. These characteristics include among others, loan performance status (default for instance), NHG uptake, loan premium, outstanding principal, interest rate, and type of loan. We define a loan as non-performing if it is either defaulted or delinquent (in arrears). In Table 1 we report descriptive statistics for these characteristics in the LLD for the years 2012 to 2018, quarter four (4), differentiating between NHG and non-NHG loans. Loan characteristics are aggregated at the household level.

3.2 Borrowers/Mortgage Characteristics

In the LLD, we speak of a higher level of aggregation when we move from loan level to borrower (or mortgage) level. Borrowers' characteristics are obtained from both the LLD and tax records. The variables from the LLD include for instance having at least one NHG loan, household income at the origination, LTV and LTI ratios and age at the origination of the oldest borrower in the household. The tax record allows us to study household current income, household wealth and even parents' wealth. In Table 1, we also report descriptive statistics for these characteristics again differentiating between borrowers that have at least one NHG loan and those who do not, as well as non-performing and underwater mortgages.

Table 1 shows high prevalence of NHG backed mortgages (27.94%)., This number is much higher when restricted to the qualified loans (65.8%); loans below the set NHG threshold. The share of mortgages underwater is 21.06%, while LTV percentage and LTI ratio average at 76.13% and 3.62, respectively. Non-performance rates are incredibly low (1.69%). It is also worth mentioning, these characteristics are much different for mortgages with NHG, non-NHG, non-performing or underwater.

Table 1: Descriptive statistics (Estimation Sample 2012 - 2018)

Descriptive Statistics	**Mean**	**Number of Observations (Household X years)**		
Share of NHG Mortgages	27.94%	2,824,713		
Share of Mortgages with NHG in New Production	65.79%	118,283		
Share of Non-performing Mortgages	1.69%	172,626		
Share of Underwater Mortgage	21.01%	2,145,962		
Share of underwater Mortgage with NHG	29.08%	821,517		
Self Employed	5.61%	573,507		
Divorced	11.75%	1,200,214		
Loan-to-Value Ratio	76.13%	10,165,147		
Loan-to-Income Ratio	3.62	10,165,147		
Loan Level Information (Annualized)	**NHG Loan**	**Non-NHG Loan**	**Non-performing**	**Underwater**
Premium	13,678	26,298	19,581	34,600
Interest Rate	3.95	3.82	4.13%	4.25%
Borrower Level Information	At least one NHG loan	No NHG loans	Non-performing	Underwater
Mean Household income at origination	48,647.62	62,766.08	51,992.4	59,752.64
Median Household Wealth	221,192.0	300,847.0	216,412.5	220,300.5
Median Property Value	192,832.0	266,329.8	211,680.0	209,267.4
Mean Outstanding Mortgage	157,789.8	193,118.4	222,089.1	255,896.1
Mean LTV at origination	90.63%	70.45%	93.05%	111.35%
Mean LTI at origination	3.70	3.55	4.45	4.55
Age of the older borrower at origination	36.65	44.91	38.87	36.96
Share Self Employed	2.79%	6.72%	7.30%	5.46%
Share Divorced/Separated	11.21%	11.94%	13.95%	11.70%
Counts	**2,824,713**	**7,284,127**	**172,626**	**2,145,962**

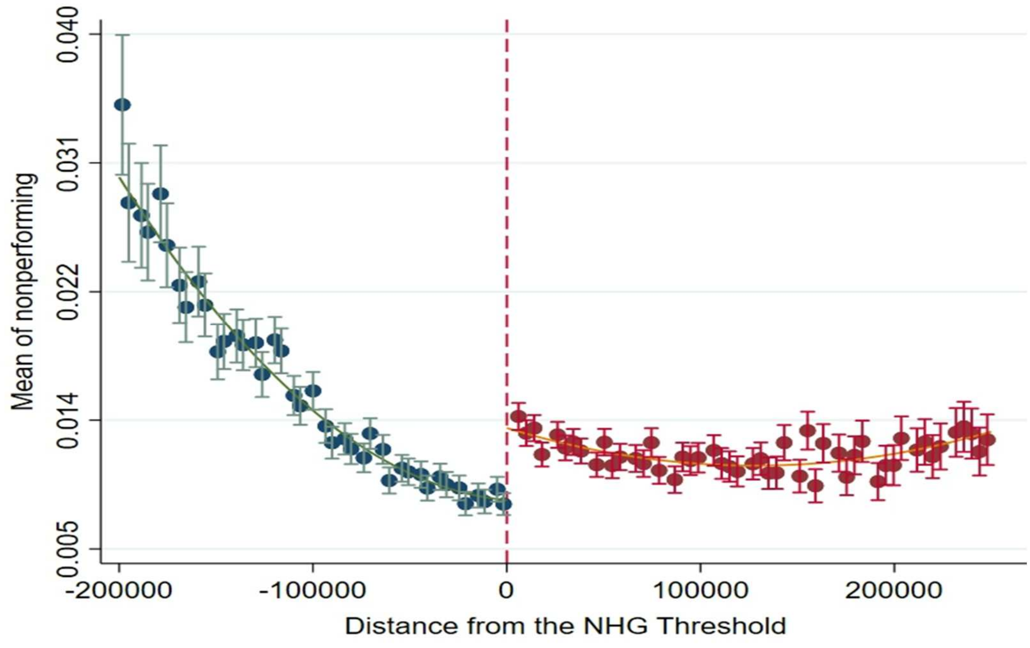
Explanatory notes: Share of NHG loans/mortgage is computed using the whole sample (total number of mortgages) rather than the qualified new loans (new productions with mortgage value below the NHG threshold). All mean differences for NHG and non-NHG borrowers are statistically different from zero at the 95% confidence level. Uptake rates of NHG are computed as a percentage of new mortgages, we assume that these are mortgages whose holders age less than 35 years. The numbers in round brackets are the counts associated with the assigned categories. All loan characteristics are aggregated at the borrower level. All differences are statistically significant at the 5% level.

3.3 Descriptive evidence

Figure 3 shows a level and slope discontinuity of non-performance around the NHG threshold, indicated by the vertical line. The horizontal axis shows the gap between the NHG threshold and the purchase price at origination. The slope on the left suggests that houses that are far below the threshold are more likely to default. A higher default among those insured could suggest a behavioural response whereby households prefer to default when covered by NHG. But an investigation of the data in the vicinity of the threshold suggests the opposite. At the threshold, the level discontinuity suggests a significantly lower non-performance among those who are insured. This is no surprise, as the NHG does not require default to reimburse residual debt. A loss incurred from selling the house can be reimbursed only if one meets the qualification criteria.

In this sense, the NHG could function not only as a tool to prevent losses but as a means to prevent defaults. However, for NHG to act as a prevention tool, the discontinuity at the threshold should be attributed to something other than default prevention. The data does not reveal a behavioural response like moral hazard, a concept often met with objections (Marshall, 1976). In some cases, such as in unemployment insurance, what is seen as moral hazard can be explained by liquidity constraints (Chetty, 2008.

Figure 3: Non-performance rate by distance from NHG threshold



Explanatory notes: The plot displays a discontinuity with evenly spaced bins, utilizing variance-mimicking spacing estimators. The NHG threshold varies over time, and the distance is calculated as the difference between the mortgage value and the NHG threshold. The mean mortgage performance rate represents the performance rate for every EUR 5000 distance from the threshold. Borrowers with values below zero are covered by NHG insurance and those above do not.

If moral hazard is not at play, we propose alternative explanations for the higher share of defaults among those with insurance. One key factor depends on institutions. The LTI rules mentioned above connect income to maximum allowable indebtedness, and consequently to house values, which ultimately determine the distance from the NHG threshold. This is because the majority of borrowers typically aim to borrow close to the maximum permitted limits for LTI and LTV. In our analysis, we need to consider income (and potential proxies for income volatility) to properly account for these factors.

Figure 4 shows this by plotting again the rate of non-performance, but this time not only relative to the distance from the NHG threshold (lower horizontal axis) but also to income (upper horizontal axis). The figure shows that there is a close vicinity in the pattern of non-performance between income and property value; as said this is due to the LTI-caps.

Examining the characteristics of the mortgage household shocks, like divorce of unemployment (as these are two of the insurance qualification criteria) is crucial. Notice that the various drivers of default suggest different policy responses. If there is moral hazard, adjusting the qualification criteria might be necessary. On the other hand, if adverse selection is prevalent, lifting or removing the threshold might be considered. But if income emerges as the main driver, refining the LTI regulations might be warranted.

Figure 4: Non-performance rate relative to the Household Income at origination (left) and to the distance from NHG threshold (lower horizontal axis) or income (upper horizontal axis).

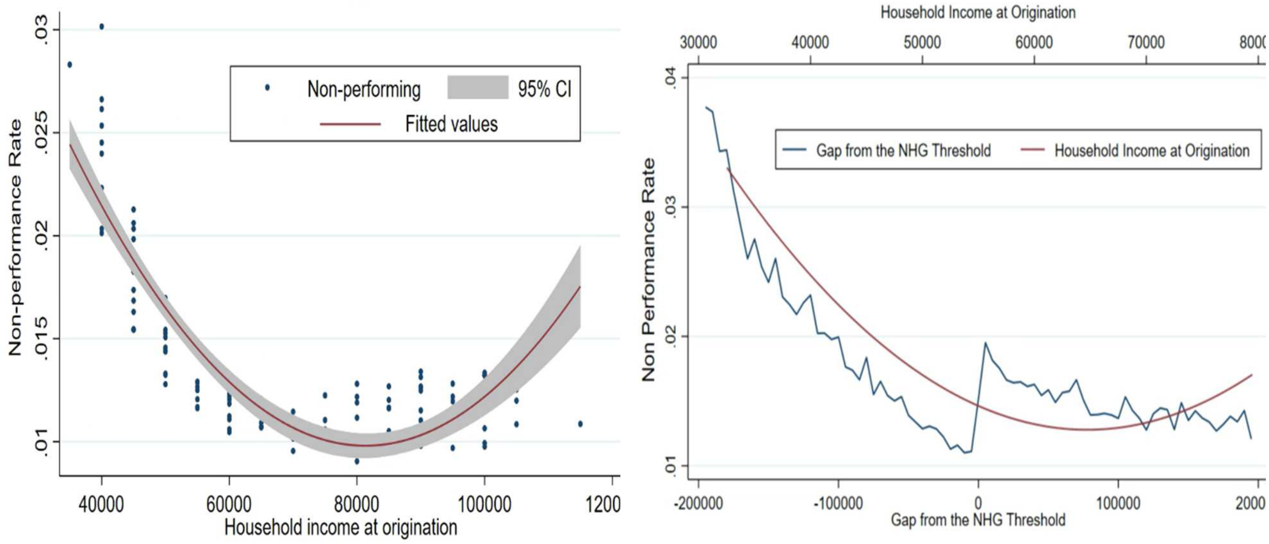


Figure 4A

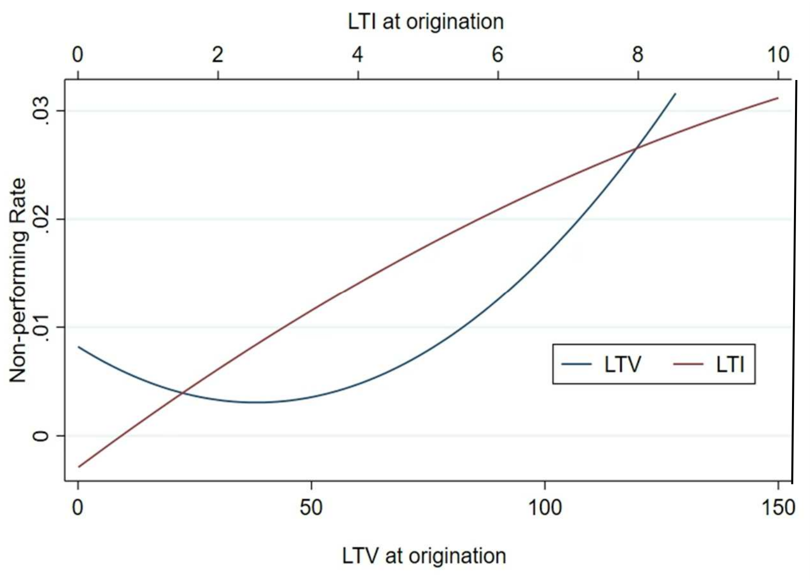
Figure 4B

Explanatory Notes: gap from the NHG threshold is the difference between the household value at origination of the mortgage and the NHG threshold for that specific year.

Beyond the eligibility criteria (divorce, unemployment, disability, death), a precondition is necessary to activate NHG, that is the existence of the residual debt or, put differently, an underwater mortgage.

In Figure 5 it is evident that there is a positive and nonlinear relation between the borrower's risk profile and the probability of default. LTV and LTI ratios commonly used to gauge creditworthiness increase with the probability of default; but while the first is convex the second is concave. A high LTI or LTV ratio implies that borrowers may more easily be borrowing more than their ability to repay, especially when considering other risk factors such as age or income shocks. The convex shape of LTV curve provides interesting insights. When LTV increases substantially, we observe an acceleration in the probability to default. This evidence coincides with the findings of Haan and Mastrogiacomo 2020. However, this pattern does not hold for the Loan-to-Income (LTI) ratio, where the probability of default somewhat slows down as income increases. One plausible explanation is the heterogeneity in LTI rules across the population, with higher income individuals having higher LTI caps each year. As those with higher income, and potentially higher LTI caps, are less likely to default, this observation aligns with the evidence presented in Figure 4B.

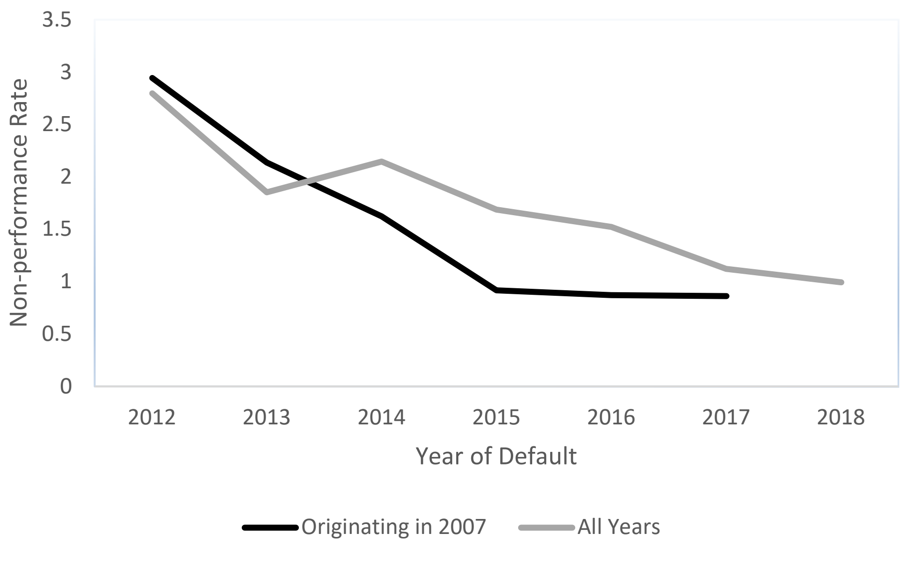
Figure 5: Non-performing Rate by LTV and LTI at the origination of the mortgage



Explanatory Note: LTI and LTV are both borrower level indicator recorded at the inception of the mortgage. Both trends are quadratic fits.

Figure 6 presents default rates by year of default. It shows a downward trend, indicating a reduction in non-performance rates of mortgage since the end of the global financial crises.

Figure 6: Non-performance Rate by year



Explanatory Notes: The dark black line is the trend of non-performance rate for all mortgage which were originated in the year 2007, while the grey line is for all mortgage independent of the year of Origination. Defaults before 2012 are not observed in our dataset.

4 Theoretical Model

In what follows, we present a theoretical framework to formally explain both how NHG works and how we can rationalize defaults using revealed preferences. We follow the literature on option value models, where the intertemporal setting allows one to not necessarily choose to default when the mortgage goes underwater.

Assume that households have a mortgage originating in $t = 0$, with a maturity period of 30 years. Each period homeowners can decide to pay the mortgage premium or to default. If they default, they become renter for the rest of the time. If they pay the premium, they face the option to default again in the next period. These options determine how home equity evolves over time in the ways specified below. Households can default if defaulting is optimal at any time during the life of the mortgage based on the value of their intertemporal utility.

Within each period s the household obtains utility from nondurable consumption and terminal home equity. The rationale for including home equity directly in the utility function is that generally consumers derive utility from home equity directly and home equity can change from property to rental dwelling as a result of defaulting.[3] From now on, the household indicator is suppressed. We define non-durable household consumption c_s as the residual income after paying mortgage premiums for mortgage owners or rent for tenants:

$$c_s \in \{c_s^m, c_s^r\} = \begin{cases} Y_s - M_s & if\ performing\ (case\ m) \\ Y_s - R_s & if \quad otherwise\ (case\ r) \end{cases} \tag{1}$$

where Y_s is the household income, M_s is the mortgage premium, and R_s is the yearly rental price. By equation 1, we abstain from modelling household savings. Home equity (terminal housing wealth, sometimes referred to as just housing wealth) h_s is set to the difference of the current value of the house and the outstanding mortgage (which is zero in the final period):

$$h_s \in \{h_s^m, h_s^r\} = CVO\left(1 + \sum_{t=0}^{s} hp_t\right) - \Phi_s, \tag{2}$$

where $\sum_{t=0}^{s} hp_t$ is cumulative house price inflation since the origination of the mortgage to time s and Φ_s the outstanding mortgage balance (residual debt) of borrower at time s, if any, while CVO_i is the collateral value of the mortgage at the origination, in this case, the originating value of the property. Since Φ_s decreases each year due to the payment of mortgage premiums, equation 2 implicitly implies that home equity for a borrower with a performing mortgage, h_s^m, is increasing with time as long as the house value does not drop by an amount greater than the mortgage premium.

For borrowers who default on their mortgage, the terminal home equity outcome h_s^r hinges on

[3] Campbell and Cocco (2015) argue that a utility function that only depends consumption relies on the assumptions that utility is separable in housing and consumption and the value of housing remains constant over time. Hatchondo et al. (2015) also include housing explicitly in their utility function as their model allows for changes in housing due to decisions taken.

whether the mortgage is insured by NHG or not. If borrowers' mortgages are insured by NHG, they lose their entire home equity when they default (if the mortgage is above water, NHG is irrelevant). However, they do not carry outstanding debt in case the house is sold with an underwater mortgage. On the other hand, borrowers without NHG experience the loss in value when they default while in addition, they may have outstanding mortgage debt if the property's selling with a mortgage that is underwater.

Summing up, different scenarios for h_s^r, defined as the value of the house minus residual debt in equation 2, can occur. If it is positive one is left with funds. If it is negative, one either incurs a loss (if one had no NHG, or if NHG did not refund) or ends up with zero (if one has NHG and qualifies for the reimbursement). In contrast with early models of default (Foster and Van Order, 1984, 1985) that include underwater status of the mortgage as a binary variable, we allow home equity to take values across the entire real line. This flexibility allows for greater variation. This is important because borrowers do not necessarily default as soon as home equity turns negative. They might prefer to wait since the default entails costs, it is irreversible and house prices might rise soon again. As the incentive to default decreases with increasing home equity, according to the option value theory, a borrower can choose the preferred action at time t given h_s.

Households' within period utility from consumption and home equity follows a pairwise separable log utility.

$$U(c_s, h_s, P, D) = \begin{cases} \log(c_s^m) + \log(h_s^m) + \varepsilon_s, & if\ performing \\ \log(c_s^r) + \log(h_s^r) + \vartheta_s, & otherwise \end{cases}, \tag{3}$$

where $c_s \in \{c_s^m, c_s^r\}$ and $h_s \in \{h_s^m, h_s^r\}$ are defined by equations 1 and 2, respectively. Notice that the values for both c_s and h_s depend on the performance of the mortgage. ε_s and ϑ_s are random parameters accounting for time-varying individual effects. Our pairwise utility function shows the difference in the utility from housing equity for a homeowner and a tenant. Stock and Wise (1988) assumed that the error components ε_s and ϑ_s follow a random walk process (as seen in the next equation). This error component may reflect unobserved characteristics of borrowers which may affect their preference differently, or other observed financial decisions which affects their choice as to whether to continue financing their mortgage or making the choice to default.

$$\varepsilon_s = \rho\varepsilon_{s-1} + \alpha_{\varepsilon s}, \quad E_{s-1}(\alpha_{\varepsilon s}) = 0, \tag{4}$$

$$\vartheta_s = \rho\vartheta_{s-1} + \alpha_{\vartheta s}, \quad E_{s-1}(\alpha_{\vartheta s}) = 0,$$

Let $V_t(D)$ be the intertemporal utility of borrowers at time t who pay their mortgage up to year $D-1$ and default at D (with $D \geq t$). Defaulted households pay the rental price for the rest of their lives. We assume that a defaulted household cannot take another mortgage or buy a house for the rest of their lives. We can write this function, given the discussion above, as:

$$V_t(D) = \sum_{s=t}^{D-1} \Delta^{s-t} U_p(c_s^m, h_s^m) + \sum_{s=D}^{S} \Delta^{s-t} U_D(c_s^r, h_s^r), \tag{5}$$

where Δ is a time preference parameter ($0 < \Delta < 1$), $U_p(c_s^m, h_s^m)$ is the utility of the borrower who continues to pay its mortgage premium and $U_D(c_s^r, h_s^r)$ is the utility of the borrower who defaults at time $D-1$ and pays rent from year D onwards. S represents the maturity period ($S = 30$) and thereby defines the relevant time horizon for the decision to default (or not). Following Stock and Wise (1998), we define $G_t(D)$ as the differential utility between defaulting in the future and defaulting now. In other words, $G_t(D)$ is the incentive function defining the incentive to keep performing. Let D^* be the optimal time of defaulting in the future, i.e. $E_t V_t(D^*) = \max_{D=t+1,\ldots,S+1} E_t V_t(D)$, where $E_t V_t(D^*)$ is the expectation of the value function at year t when one defaults, given the household pays the mortgage premium up to year $D-1$ while $E_t V_t(t)$ is the value of defaulting in t. $D = S+1$ represents the case of never defaulting.

$$G_t(D^*) = E_t V_t(D^*) - E_t V_t(t), \tag{6}$$

For $G_t(D^*) < 0$, the borrower defaults in t, as the expected value of default in t is larger than the value for not defaulting. If $G_t(D^*) \geq 0$, one does not default, and repeats the decision process in $t+1$.

In our sample, each household is observed for at most $\tau = 1, \ldots, 11$ years.[4] If the household is observed in year $t \leq 11$, the option value model assigns the probability

$$P(default\ in\ t) = \prod_{\tau=1}^{t-1} P(G_\tau(D_\tau^*) > 0) P(G_t(D_t^*) \leq 0)$$

representing the joint probability of not defaulting in $\tau = 1, \ldots, t-1$ and defaulting in *t*.

For a household that does not default during the observed 11 years we get

$$P(no\ default) = \prod_{\tau=1}^{11} P(G_\tau(D_\tau^*) > 0)$$

In the data we observe the value of chosen alternative D, and must compute the value of the counterfactual alternative. This means for instance that if we observe someone who never defaults, we must assign the value of defaulting in each possible period. Pivotal to this is imputation of the cost of housing in the alternative state. As we do not observe rents, we impute them using maximum rent affordability caps, and perform sensitivity analysis to this parameter.

Substituting equations 3 and 4 into equation 5, we can decompose the borrower defaulting

[4] We follow a sample of mortgages all originating in 2007 with a 30-year maturity period. Given the availability of the Loan level data only from 2012 to 2018, we observe the mortgage from 2007 (retrospective) to 2018. This gives us 11 years of observation.

decision function.

$$G_t(D) = E_t \sum_{s=t}^{D-1} \Delta^{s-t} (\log(c_s^m) + \log(h_s^m) + \varepsilon_s) + E_t \sum_{s=D}^{S} \Delta^{s-t} (\log(c_s^r) + \log(h_s^r) + \vartheta_s) - E_t \sum_{s=t}^{S} \Delta^{s-t} (\log(c_s^r) + \log(h_s^r) + \vartheta_s), \quad (7)$$

$G_t(D)$ can be decomposed into three components,

$$G_t(D) = C_t(D) + H_t(D) + Z_t(D), \quad (8)$$

Where $C_t(D)$ is the difference in lifetime consumption at time t, for the mortgage holder who continues to pay their mortgage premium up to the optimal time to default and pay a rent their after, $H_t(D)$ if the home equity; the difference in terminal housing wealth for the option to perform on the mortgage and that of defaulting. $Z_t(D)$, is a random effect containing ε_s and ϑ_s.

Stock and Wise (1998) estimated the parameters of this model with maximum likelihood, while Samwick and Wise (2003) specified the parameters and estimated the outcomes in a probit regression. Instead, we estimate the value of the differential utilities in equation 8 (hence equation 7), and use the terms as regressors in a conditional logit regression where the outcome variable is performing. The last term ($Z_t(D)$) in equation 8 is assumed to be zero for simplicity. To compute the differential utility, it is assumed that the observed choice D is preferred, among all 11 alternatives (j can be equal to never defaulting or defaulting in the years between 2008 and 2017), we also obtain the value as if the borrower was to choose the alternative option and compute the difference.

Ideally, we would estimate the $E[maxV_t(D)]$ and to do so, a typical approach is to use the nested fix point algorithm of Rust (2000). However, this method is based on backward induction, and this implies observing loans till maturity. Estimation of this model based on the selected sample of mortgages defaulting between 2007-2018 delivered, as expected a positive marginal utility for consumption and housing wealth (with coefficients 2.853 and 0.0174, respectively). However, if we want to use all mortgages observed between 2007-2018, we estimate the option value version of the model. This implies computing first $V_t(D)$, then $E_t V_t(D)$ and finally the $\max[E_t V_t(D)]$.

4.1 Empirical Implementation

In an option value model, individuals choose whether to default or not every year, considering in each year the future value of each present option. This is different from a dynamic programming situation, where one makes choices looking at all possible time paths during the life-cycle. As we only have eleven periods and choices (between performing or not), we will present only the results of an option value model where intertemporal choices are made each year

From equation 3, individuals are assumed to have separable log utilities in present (c_s) and future consumption (h_s, home equity) and the error term on the estimated utility is assumed to be a logistic error. Consider the choice made at time t, with $t \in \{1, \ldots, 11\}$. Denote $\bar{V}_t(D) = E_t V_t(D)$ (expectation without the logistic error part). Defaulting at t means $D = t$ gives the maximum utility so with logistic errors we get the probability to default at t (out of any other option ranging from $D = t + 1, \ldots\ldots, S$. Notice we drop the individual borrower identifier, i.

$$P(D = t) = \frac{\exp(\bar{V}_t(t))}{\sum_{s=t}^{30} \exp(\bar{V}_t(s))} = \frac{1}{1 + \sum_{s=t+1}^{30} \exp(G_t(s))} \tag{9}$$

The probability to not default at t then is $P(not\ default\ at\ t) = 1 - P(D = t) = \frac{\sum_{s=t+1}^{30} \exp(G_t(s))}{1+\sum_{s=t+1}^{30} \exp(G_t(s))}$.

All regressors are prepared separately, outside the likelihood function, by predicting income, rent, non-durable consumption, and home equity. Household income is observed at inception of the mortgage and assumed to grow at a random positive rate of at most three (3%) percent, while the rental price is set at 25% of household income (with sensitivity checks for rates between 22 – 38% of income).

Let X be the vector of the elements in the utility function and their interaction with observables, such as age and the ownership of NHG, then the probability that option j is chosen can be written as:

$$Pr\left(D | X_{it}\beta_j + \alpha_i\right) = \frac{exp(X_{it}\beta_D + \alpha_i)}{\sum_{s=t}^{S} exp(X_{it}\beta_D + \alpha_i)} \tag{10}$$

Where, α_i is normally distributed indicator for unobserved heterogeneity. As the choice probabilities are conditioned on the unobserved heterogeneity, α_i, the likelihood contribution can be obtained by integrating over the distribution of unobserved heterogeneity. The likelihood function for a multinomial logit with a random intercept boils down to:

$$L = \sum_{i=1}^{N} \sum_{t=1}^{S} \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left(\frac{exp(X_{it}\beta_D + \alpha_i)}{\sum_{s=t}^{S} exp(X_{it}\beta_D + \alpha_i)} \right) d_{ijt} f(a) da \tag{11}$$

Where d_{ijt} is a dummy which is equals to 1 if mortgage holder i choose option j at time t. For identification, we define a reference category. In addition S is year 2018 (so the second summation looks at all periods observed in the data) while T refers to the planning horizon of the individual in each of these periods and thus embodies the 30th year of the mortgage. As explained we simulate the development of the main utility parameters for each individual in each year, using assumptions on how these elements develop to maturity. The approach we follow is to maximize a simulated likelihood. We draw R times values of a from its distribution, compute the likelihood, and average it out over the R draws. So, we don't estimate the exact likelihood but a simulated one, namely:

$$L = \sum_{i=1}^{N} \frac{1}{R} \sum_{r=1}^{R} \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left(\frac{exp(X_{it}\beta_D + \alpha_i)}{\sum_{s=t}^{S} exp(X_{it}\beta_D + \alpha_i)} \right) d_{ijt} \tag{12}$$

Estimates of this model are presented in Table 2. Our model delivers, as suggested by theory, also a positive marginal utility for consumption and home equity, implying both a desire for higher consumption and higher terminal home equity. An increase in the utility of consumption is associated with a lower probability of default. In a conditional logit setting, parameters can be estimated only when they vary with the j alternatives, as it is the case for consumption and home equity here. Including alternative-constant variables (such as age or the NHG status) is only possible when these terms interact with alternative-varying utility parameters. Multiplying either consumption or home equity by these fixed variables is enough to affect the marginal rate of substitution. We chose to report results for the interaction of these alternative-constant variables with consumption. The interaction between consumption and NHG shows that marginal utility of consumption increases when one has NHG. Also, the results show that marginal utility decreases with age. Finally, the table also shows that unobserved heterogeneity is not statistically significant.

Table 2: Estimates of option value model

	Full Model	Age Disaggregation		Employment		Type of House		Location	
		Young (Age 35 or younger)	Old (Older than 35)	Employees	Self-employed & others	Detached	Apartments	Randstad	Other cities
Consumption	2.853***	19.57***	1.108***	3.797***	0.695***	2.893***	2.886***	2.857***	2.944***
	(0.0266)	(0.248)	(0.0326)	(0.0341)	(0.0361)	(0.0310)	(0.0512)	(0.0663)	(0.0288)
Home Equity	0.0174***	-0.17***	0.229***	-0.017***	1.100***	0.065***	-0.0747***	-0.087***	0.0368***
	(0.000740)	(0.0015)	(0.0014)	(0.00079)	(0.00617)	(0.0009)	(0.00134)	(0.00193)	(0.000815)
Age	-0.0099***	-0.49***	0.017***	-0.024***	0.003***	-0.012***	-0.00860***	-0.026***	-0.00990***
	(0.000534)	(0.0076)	(0.0006)	(0.0007)	(0.0007)	(0.0006)	(0.00103)	(0.00136)	(0.000576)
NHG	3.906***	1.962***	4.650***	3.262***	9.194***	4.881***	1.593***	1.365***	4.310***
	(0.0264)	(0.0544)	(0.0311)	(0.0276)	(0.109)	(0.0329)	(0.0440)	(0.0579)	(0.0292)
a	-21.75								
	(702.9)								
Observations	4,805,933	1,031,592	3,774,341	3,830,937	974,996	3,689,279	1,116,654	507,375	4,298,558

Explanatory notes: Standard error in parentheses. Age and NHG are interacted with log consumption (the differential utility from lifetime consumption). The model is estimated by a conditional logistics regression, with fixed effects for the year of default and cohort effects. Model standard errors are VCE clustered to the ID level. Estimates are from an unbalanced panel; borrowers drop soon as the default. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Heterogeneity Analysis

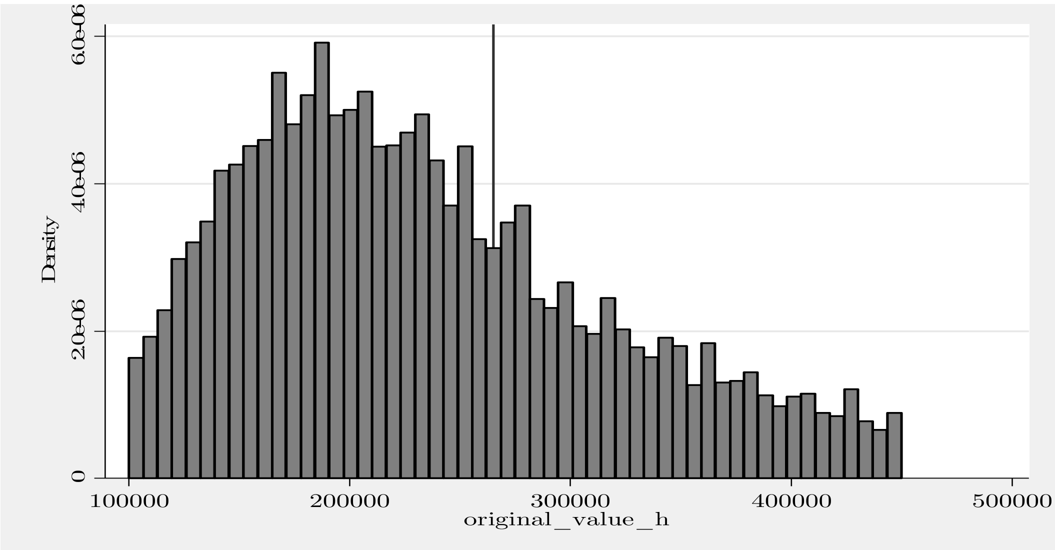
To relate the estimated coefficients to observed heterogeneity, we estimate the model above for some subgroups (Table 2). We show heterogeneity for age, employment status, type of house and the location of the city. The main result of marginal utility being positive and increasing for NHG participants remains for all subsamples. The coefficient associated to the marginal utility of home equity is instead negative for young, employees, who have an apartment in the most metropolitan area of the Netherlands (the Randstad Area that is the cities of Amsterdam, Utrecht, Rotterdam, or the Hague). A possible interpretation of this is that this group might have a preference for renting rather than owning at this stage of their life.

Above we discussed that a crucial step in the estimation of these parameters depends on the value of non-chosen alternatives. In all cases this imposes to impute the cost of renting rather than paying a mortgage premium. In the computations above, we assume that those who default rent an apartment that will cost 25% of their gross income (we call this percentage the 'rental rate'). This is in line with affordability caps that are also used to assess the share of income that can service mortgage debt. This means that the rental rate is quite low by current standards, as tenants for instance do not need to pay maintenance. In Appendix 1, we test the sensitivity to this assumption, by reducing and (most realistically) increasing the rental rate between 22% and 38%. Notice that already at 35% most stress test models assume households as being defaulted.
Our results are not sensitive to an increase in the rental rate, a further decrease makes renting even cheaper, and mortgage holder prefer renting to continuing to perform on their mortgage.

5 Quasi Natural Experiment

To confirm the results presented above, we also present the results of a quasi-natural experiment based on the discontinuity implied by the institutional design of NHG. We present a reduced-form regression discontinuity model to study the effect of having NHG insurance on the probability of default. This is a legitimate approach as the NHG threshold qualifies as a sharp cutoff point (c) that allows assignment to treatment. Again, our outcome variable is the binary indicator of default (Y_{D_i}), the assignment variable is the value of the house at purchase (h_{i1}) and the treatment status implies that $NHG = 1[h_1 < c]$. In this design two assumptions are relevant, namely that the treatment effect is constant (it does not differ by household), and we assume continuity. This means that unobservable v_i are such that $E(v_i|h_{i1} = h_1)$ is continuous in h_1. First, we try to rule out adverse selection. Pre-selection in the form of adverse selection should interfere with the anticipated effect of NHG. According to Jullien (2000), this implies bouncing around the NHG threshold. In Figure 7 we notice no bouncing around the threshold but a continuity of units that suggests that borrowers do not strategically position themselves below the threshold.

Figure 7: Distribution of borrowers around the NHG threshold in 2014



Explanatory Notes: This is a density plot with density on the vertical axis and value of the property at origination on the horizontal axis. The black vertical line is the NHG threshold as at 2014, the origination date of the mortgage. The figure includes only mortgages originated in 2014. Source; LLD and own computations

Defaults also reduce borrowers' home equity (Mian and Sufi, 2011); therefore, we relate non-performance to the ability to pay, liquidity of the borrower(s) and/or affordability of the mortgage, (Vandell (1995), Ambrose et al. (1997)). We evaluate whether non-performance depends both on the discontinuity caused by the NHG qualification or it can be explained solely by liquidity (household wealth and income) and mortgage affordability indicated by specific loan characteristics. When a borrower experiences negative home equity, defaulting with NHG reimbursements is beneficial, but not strictly necessary as selling the house and recognizing the loss is enough to initiate an insurance claim.

So we describe the effect of insurance on the probability of not performing using the following regression discontinuity design:

$$Y_{D_{it}} = \alpha + \beta_1 uw_{i,t} + \beta_2 NHG_i + \beta_3 NHG_i(uw_{i,t}) + v_{i,t}, \qquad (13)$$

$Y_{D_{it}}$ is a binary variable, for each borrower i, it takes the value 1 if the borrower is not performing and zero otherwise. $\alpha, \beta_1, \beta_2,$ *and* β_3 are parameters, where β_1 measures the effect of having negative home equity. β_1 is expected to be positive, as the underwater status in an incentive to default. β_2 measures the impact of NHG on the probability of default, and β_3 the differential effect of NHG conditional on the mortgage being underwater. In this model β_3 is the interesting parameter, as the insurance can be claimed only if one has residual debt, and this only happens if the mortgage is underwater. Estimates of this model are shown in Table 3 where both the parameter on NHG and its interaction with underwater status are negative and statistically different from zero. There we also report two more estimates of β_3 where we reduce the sample to the observations that are closer to the threshold and two more specifications dropping those far away from the threshold. Results are consistent across these checks, meaning that these are not only locally valid.

Table 3: Regression results (Restricted Samples)

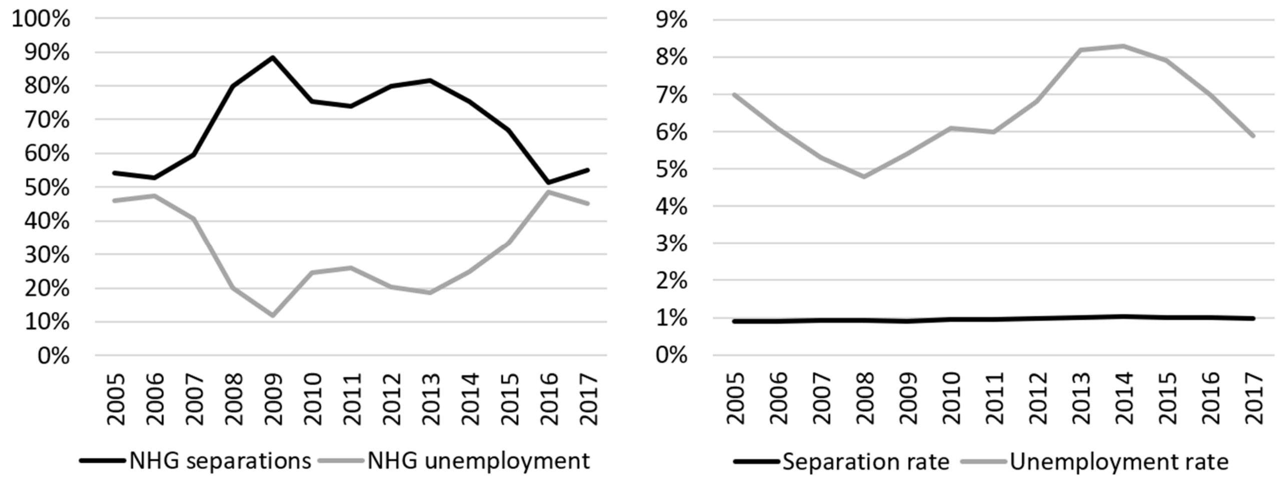
	All	Drop 5% tails	Drop 10% tails	Drop at threshold	Drop 5000 EUR from threshold
Underwater (β1)	0.0355***	0.0351***	0.0351***	0.0355***	0.0353***
NHG (β2)	-0.0005*	-0.0006**	-0.0009***	-0.0005*	-0.0004
NHG * underwater (β3)	-0.0179***	-0.0211***	-0.0215***	-0.0178***	-0.0173***
Constant	0.0129***	0.0131***	0.0130***	0.0129***	0.0129***
Observations	1,594,766	1,261,260	1,120,874	1,569,792	1,522,232
R-squared	0.0090	0.0091	0.0094	0.0090	0.0089

Explanatory Notes: All models are linear probability. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. NHG, Unemployment and divorce are dummies with value 1 if the borrower has NHG, is unemployed or divorced. The dependent variable, Not performing, is a binary variable with value one (1) if the borrower chooses to default or if the borrower is in delinquency and zero if they continue paying their mortgage or resale. NHG takes value one (1) if the borrower has NHG insurance implying that their house value is less than the NHG threshold. Data use for only 2014.

These baseline estimates support the observation that having insurance reduces the likelihood defaults. It's important to note that filing a claim requires meeting certain conditions. Data from the insurance fund reveals a significant rise in claims related to separations during the asset prices crisis. This trend reversed once the crisis subsided. Interestingly, despite this data, national statistics indicate no notable increase in separations during that crisis period. (see Figure 8).

To investigate further, we need to closely examine whether separations among those with insurance led to more defaults. To do this, we expand the RDD by including controls for separations. These controls involve a dummy variable called "divorced" which signals separations, and its interaction with the NHG indicator.

Figure 8: Honoured claims of NHG participants by qualification criteria (left), versus national separation and unemployment rate (right)



Explanatory notes: Separations and unemployment within NHG are expressed as a percentage of the total. Source: NHG yearly reports, different years. The separation (unemployment) rate conveys the number of separations (unemployed) every 100 inhabitants. Source: CBS Statline.

$$Y_{D_{it}} = \alpha + \beta_1 uw_{i,t} + \beta_2 NHG_i + \beta_3 NHG_i\left(uw_{i,t}\right) + \beta_4 NHG_i\left(separated_{i,t}\right) + \beta_5 separate_{i,t} + \beta_6 unemployed_{i,t} + \beta_7 NHG_i * unemployed_{i,t} + \delta X_{i,t} + v_{i,t}, \tag{14}$$

Put differently, we want to check how separations and NHG affect defaults. Separating is a choice within the household. Other criteria like disability, death, and unemployment are usually involuntary. If no effect of insurance, like moral hazard, plays a role, we would expect no significant difference in defaults between those separating or not with and without insurance. Instead, our findings show a positive and significant impact of insurance on defaults ($\beta_4 > 0$), even after considering other factors (see Table 4). These covariates are especially relevant given the slope discontinuity in Figure 3. The inclusion of the right set of covariates will also help capture the effect of the mortgage value and NHG on non-performance away from the threshold (Angrist and Rokkanen, 2015).

In Table 4, we show the impact of NHG insurance and its interactions with underwater status and eligibility criteria (separations and unemployment) on the probability of non-performance rate in four different specifications. These specifications differ because we add the other covariates. Even when adding these additional variables, the effect of the main parameters stays similar.

Contrary to the setup of traditional stress test models, the relative value of home equity (above captured by the underwater dummy and the LTV variable) shows a significant and positive effect on the probability of not performing.

Further the borrowers' income, wealth, and the affordability of the mortgage (the DSTI ratio), are negatively related to non-performance.

Table 4: Regression Results

	2014		2012-2018	
	Model A	Model B	Model C	Model D
Underwater (β1)	0.0331***	0.0238***	0.0293***	0.0223***
NHG (β2)	-0.0021***	-0.0058***	-0.0024***	-0.0056***
NHG * underwater (β3)	-0.0164***	-0.0141***	-0.0157***	-0.0140***
Separated (β4)	0.0044***	-0.0006	0.0042***	0.0003
Separated*NHG (β5)	0.0006	0.0016*	0.0007*	0.0016***
Unemployed (β6)	0.0165***	0.0119***	0.0129***	0.0097***
Unemployed * NHG (β7)	0.0101***	0.0102***	0.0075***	0.0077***
Prepaid (Cum. Prep. >0)				0.0031***
Prepaid * NHG				-0.0028***
Debt service to income ratio		0.0032***		0.0016***
Income (ihs)		-0.0128***		-0.0099***
Assets (ihs)		-0.0033***		-0.0027***
LTV at origination		0.0181***		0.0121***
Current age		-0.0003***		-0.0006***
Self-employed at origination		0.0073***		0.0058***
Dummies for year of origination	yes	yes	yes	yes
Year of birth dummies	yes	yes	yes	yes
Constant	0.0513***	0.2283***	0.0278***	0.1737***
Observations	1,586,666	1,586,666	10,108,840	10,108,840
R-squared	0.0110	0.0157	0.0093	0.0128

Explanatory Notes: All models are linear probability. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. NHG, Unemployment and divorce are dummies with value 1 if the borrower has NHG, is unemployed or divorced. The dependent variable, Not performing, is a binary variable with value one (1) if the borrower chooses to default or if the borrower is in delinquency and zero if they continue paying their mortgage or resale. NHG takes value one (1) if the borrower has NHG insurance implying that their house value is less than the NHG threshold. Household income and wealth are the gross income and wealth of all household members, respectively.

The probability of default decreases as the borrower's age increases. This observation aligns with the equity theory of default proposed by Jackson and Kaserman (1980), which suggests a connection between mortgage maturity and default risk. Mortgages nearing maturity tend to have lower default risk, likely due to higher home equity (but also higher income insurance of the borrower and survival bias). Younger borrowers, being closer to the start of their mortgage, typically have less home equity and are farther from maturity. Despite having lower loss given default, especially with NHG insurance protection, younger borrowers may be more susceptible to economic shocks, making them inherently riskier. As self-employed, although are typically older, are more vulnerable to shocks, their more volatile income could explain the positive relation to defaults.

So far, we have not discussed the relative magnitudes of the effects presented. In Appendix 3, we do so by showing marginal effects. We see NHG does not affect much the probability of default, somewhat larger is the effect of separations. Unemployment and underwater status display larger and similar effects in terms of magnitude.

6 Robustness Checks

In Table 5, we report some robustness checks and placebo tests for our baseline result in Model B (see Table 4). As distance from the NHG threshold is a potential concern, in the upper panel of the table we drop first the 5% and 10% tails of the distribution (checks 1 and 2). We also remove from the whole sample the observations near the threshold (checks 3, 4 and 5).[5] All these checks reveal that the results stay close to the baseline estimates, confirming their robustness.

In the lower panel, we conduct several placebo tests. Here, we assign treatment based not on NHG participation but on borrowers individual income or wealth. This allows us to assess whether NHG participation merely acts as a signal for these two variables, thus challenging the assumption of treatment assignment being exogenous. In all scenarios, the estimates switch signs and become positive, which goes against the fundamental evidence that insurance has decreased (rather than increased) defaults. This suggests that neither income nor wealth are suitable variables for explaining the default patterns observed in the data, while NHG participation is.

Table 5: Robustness and placebo check for model estimates (2014)

Model	NHG*underwater (β3)
Baseline	-0.0141***
Robustness checks:	
Dropping 5% tail	-0.0170***
Dropping 10% tail	-0.0175***
Drop observation €5000 from threshold	-0.0134***
Drop observation €10000 from threshold	-0.0123***
Observations within €50000 from threshold	-0.0165***
Placebo tests:	
Treatment: income above €40000	0.0370***
Treatment: income above €60000	0.0130***
Treatment: fin. wealth above €50000	0.0041
Treatment: fin. wealth above €100000	0.0277***

In Table 6, we explore various policy alternatives using the reduced form model (equation 10). First, we examine the prospect of raising the NHG threshold by increments of 10, 50, or 100 thousand euros. This would potentially extend insurance coverage to properties above the average purchase prices. The resulting reduction in default rates would be at most 14% (calculated as (2.04% - 1.76%) / 2.04%). However, it s important to note that this policy adjustment might entail increased costs for the NHG fund.

Second, we explore the option of raising the limits on Loan-to-Income (LTI) and Loan-to-Value (LTV) ratios. The Dutch National Bank (DNB) reports that over 50% of borrowers utilize more than 90% of their borrowing capacity, indicating the presence of binding credit constraints.[6] Although these caps are relatively high compared to international standards, policymakers often

[5] In administrative data where assignment to treatment is observed, this check is probably less interesting, relative to cases where the assignment is more fuzzy.

[6] See Figure 4 in https://www.dnb.nl/media/faxpn0vj/ofs_najaar_2020.pdf

consider adjustments to increase them. If an expansion of these limits (by 10%, 20%, or 30%) leads to a proportional rise in debt levels, it could potentially result in an increase in default rates of up to 21% (calculated as (2.47% - 2.04%) / 2.04%).

Table 6: Policy simulations for share of defaults based on 2014 data.

	Predicted defaults	St. Dev.
Baseline model	2.04%	0.0183
NHG threshold increases by 10000 euro	1.86%	0.0183
NHG threshold increases by 50000 euro	1.80%	0.0182
NHG threshold increases by 100000 euro	1.76%	0.0182
LTI limit increases by 10%	2.19%	0.0181
LTI limit increases by 20%	2.23%	0.0185
LTI limit increases by 30%	2.47%	0.0189
LTV at origination increase by 10%	2.17%	0.0187
LTV at origination increase by 20%	2.31%	0.0192
LTV at origination increase by 30%	2.47%	0.0197

6 Summary and Conclusions

This paper studies the impact of NHG insurance on non-performance rates of mortgages using Loan Level Data from the Dutch National Bank merged with current data on the household income wealth and other administrative registry from the CBS. We estimated a structural model that considers the possibilities to either stop making mortgage payments at an optimal time, thus renouncing home ownership and becoming a tenant, or to continue being performing. We then tested the results of the structural model using a reduced form approach based on a quasi-natural experiment. This revolves around the discontinuity in the assignment of NHG insurance to identify its impact of mortgage defaults.

The estimates of the structural model show that having NHG increases marginal utility of consumption. The reduced form models confirm this, as NHG appears to make borrowers less likely to default, even with the presence of an underwater mortgage. However, conditional on divorced, the probability to default increase with presence of NHG insurance. Further, we show that the downward slope before the NHG threshold is explained by conventional determinants of mortgage default: liquidity of the borrower and affordability of the mortgage.

Behind this baseline result, we also find substantial heterogeneity, indicating for instance higher marginal utility of consumption for younger borrowers. In the reduced form model instead, we see that the chance of default increases when the mortgage is underwater, when there is a divorce or separation, or if the borrower becomes unemployed. Within sample policy simulations suggest that an increase in borrowing limits by 30% would induce an increase in the default rate by 21%. While

NHG insurance does not make default more likely, borrowers with NHG are more likely to default if they go through a divorce. A within sample policy simulation for the extension of the NHG threshold by 100K suggests that the default rate would decrease by about 14%, with a potential increase for the costs of the NHG fund.

7 References

Ambrose, Brent W, Richard J Buttimer Jr, and Charles A Capone (1997). "Pricing mortgage default and foreclosure delay." In: *Journal of Money, Credit, and Banking*, pp. 314–325.

Angrist, Joshua D and Miikka Rokkanen (2015). "Wanna get away? Regression discontinuity estimation of exam school effects away from the cut-off." In: *Journal of the American Statistical Association* 110(512), pp. 1331–1344.

Chetty, Raj (2008). "Moral hazard versus liquidity and optimal unemployment insurance." In: *Journal of Political Economy* 116(2), pp. 173–234.

Chiappori, Pierre-André and Bernard Salanie (2000). "Testing for asymmetric information in insurance markets." In: *Journal of Political Economy* 108(1), pp. 56–78.

Foster, Chester, and Robert Van Order (1984). "An option-based model of mortgage default." In: Housing Fin. Rev. 3, p. 351.

Foster, Chester, and Robert Van Order (1985). "FHA terminations: A prelude to rational mortgage pricing." In: *Real Estate Economics* 13(3), pp. 273–291.

Geanakoplos, John (2010). "The leverage cycles." In: *NBER macroeconomics annual* 24(1), pp. 1– 66.

Gerardi, Kristopher, Kyle F Herkenhoff, Lee E Ohanian, and Paul S Willen (2018). "Can't pay or won't pay? Unemployment, negative equity, and strategic default." In: *The Review of Financial Studies* 31(3), pp. 1098–1131.

Gerlach-Kristen, Petra, and Seán Lyons (2018). "Determinants of mortgage arrears in Europe: evidence from household microdata." In: *International Journal of Housing Policy* 18(4), pp. 545– 567.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2013). "The determinants of attitudes toward strategic default on mortgages." In: *The Journal of Finance* 68(4), pp. 1473–1515.

Haan, Leo de and Mauro Mastrogiacomo (2020). "Loan to value caps and government-backed mortgage insurance: Loan-level evidence from Dutch residential mortgages." In: *De Economist* 168(4), pp. 453–473.

Hallissey, Niamh, Robert Kelly, and Terry O'Malley (2011). "Of Property Lending at Irish banks."

Hatchondo, Juan Carlos, Leonardo Martinez, and Juan M Sánchez (2015). "Mortgage defaults." In: *Journal of Monetary Economics* 76, pp. 173–190.

Jackson, Jerry R and David L Kaserman (1980). "Default risk on home mortgage loans: a test of competing hypotheses." In: *Journal of risk and insurance*, pp. 678–690.

Kim, Yeorim, Mauro Mastrogiacomo, Stefan Hochguertel, and Hans Bloemen (2022). "Till debt do us part: strategic divorces and a test of moral hazard."

Kim, Y., Mastrogiacomo, M., Hochguertel, S., & Bloemen, H. (2019). Is additional indebtedness the way to increase mortgage-default insurance coverage. *Netspar Design Paper*, *115*, 1-43.

Li, Wenli and Florian Oswald (2017). "Recourse and residential mortgages: The case of Nevada." In: *Journal of Urban Economics* 101, pp. 1–13.

Marshall, John M (1976). "Moral hazard." In: *The American Economic Review* 66(5), pp. 880–890.

Mayer, Christopher, Edward Morrison, Tomasz Piskorski, and Arpit Gupta (2014). "Mortgage modification and strategic behaviour: Evidence from a legal settlement with countrywide." In: *American Economic Review* 104(9), pp. 2830–57.

Mian, Atif and Amir Sufi (2011). "House prices, home equity-based borrowing, and the US household leverage crisis." In: *American Economic Review* 101(5), pp. 2132–56.

Riley, Sarah F (2013). "Strategic default behaviour and attitudes among low-income homeowners." In: *Real Estate Finance* 29(5), pp. 22–36.

Teye, Alfred Larm, Jan de Haan, Marja Geessiena Elsinga, Francis Kwesi Bondinuba, and Job Taiwo Gbadegesin (2017). "Risks in homeownership: a perspective on The Netherlands." In: *International Journal of Housing Markets and Analysis*.

Vandell, Kerry D (1995). "How ruthless is mortgage default? A review and synthesis of the evidence." In: *Journal of Housing Research*, pp. 245–264.

Wong, T-Ch, Tom Fong, Kafai Li, and Henry Choi (2011). "Loan-to-value ratio as a macroprudential tool-Hong Kong's experience and cross-country evidence." In: *Systemic Risk, Basel III, Financial Stability and Regulation*.

Appendix

Appendix 1: Rental Rates

	Rent as a percentage of total household income					
	22%	25%	28%	30%	34%	38%
Consumption	-0.679***	2.853***	2.870***	3.723***	4.008***	3.218***
	(0.00753)	(0.0266)	(0.0264)	(0.0251)	(0.0215)	(0.0181)
Home Equity	0.0170***	0.0174***	0.0165***	0.0204***	0.0231***	0.0259***
	(0.000676)	(0.000740)	(0.000743)	(0.000760)	(0.000776)	(0.000784)
Age	0.0261***	-0.00988***	-0.00900***	-0.0251***	-0.0348***	-0.0261***
	(0.000224)	(0.000534)	(0.000531)	(0.000496)	(0.000420)	(0.000358)
NHG	5.042***	3.906***	3.559***	3.235***	2.448***	2.032***
	(0.0273)	(0.0264)	(0.0253)	(0.0240)	(0.0200)	(0.0167)

Appendix 2: Detailed Regression results with model improvements

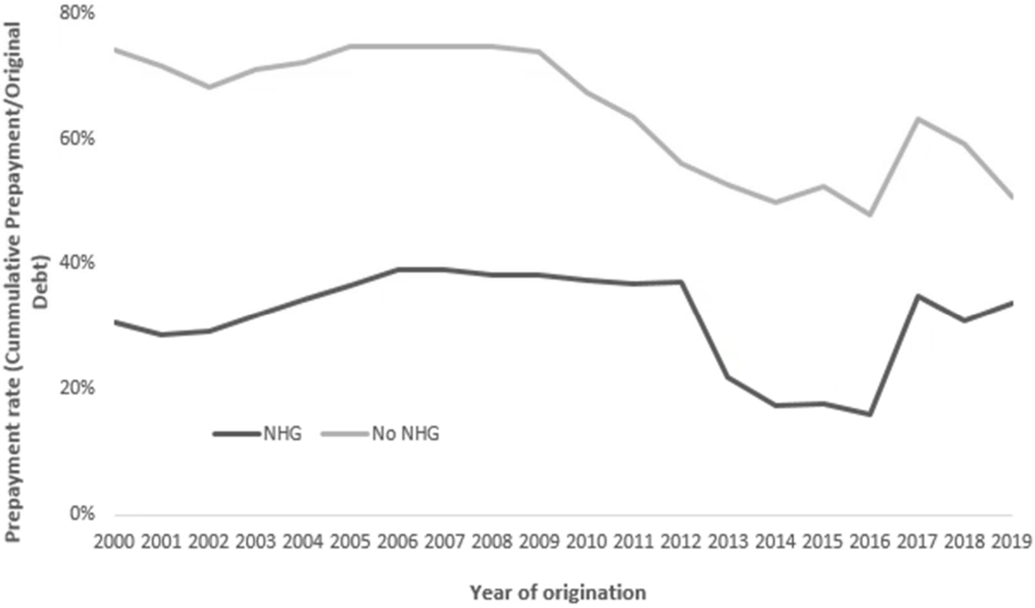
	Model 1	Model 2	Model 3	Model 4
NHG (β_2)	-0.0023***	-0.0043***	-0.0055***	-0.0054***
Underwater (β_1)	0.0293***	0.0279***	0.0232***	0.0223***
Underwater * NHG (β_3)	-0.0157***	-0.0154***	-0.0140***	-0.0140***
Divorced (β_5)	0.0042***	0.0009***	0.0002	0.0003
Divorced * NHG (β_4)	0.0008*	0.0012***	0.0013***	0.0016***
Unemployed	0.0147***	0.0115***	0.0114***	0.0116***
Debt Service to Income Ratio		0.0019***	0.0016***	0.0016***
Household Income		-0.0091***	-0.0098***	-0.0099***
Household Wealth		-0.0028***	-0.0027***	-0.0027***
LTV at origination			0.0121***	0.00121***
Age of the Older borrower				-0.0006***
Self-employed				0.0058***
Constant	0.0277***	0.1643***	0.1588***	0.1736***
Year of origination	Yes	Yes	Yes	Yes
Cohort Effect	Yes	Yes	Yes	Yes
Number of observations	10,108,840	10,108,840	10,108,840	10,108,840

Appendix 3: Average Marginal Effects

Variable	Model 1	Model 2	Model 3	Model 4
NHG	-0.00013	-0.00125	-0.00218	-0.00242
Underwater	0.01985	0.01918	0.01467	0.01471
Divorced	0.00476	0.00308	0.00269	0.00284
Unemployed	0.0120	0.0104	0.0101	0.0106

Explanatory Notes: These marginal effects are all average marginal effects and are extracted from the regressions in Table , Models 1, 2, 3,& 4 are as specified.

Appendix 4: Prepayment rate by NHG Status and year of origination of the mortgage



DeNederlandscheBank
EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl