Bank Size and Wholesale Funding Supply

Yilun Allen Shen *

November 2026

Abstract

When a depository bank loses net worth, does it contract its own lending or the funding it provides to nonbank lenders? In U.S. Call Reports, mortgage-backed and other asset-backed securities account for roughly 6% of assets for smaller banks and nearly 15% in for largest banks. I develop a model in which a wholesale claim funds a nonbank lender and serves as a private safe asset for the depository bank that holds it. Small banks hold just enough claims to meet liquidity requirements and cut both forms of lending after a loss. In contrast, large banks absorb a marginal loss of balance-sheet capacity entirely by cutting wholesale funding. In a U.S. banking industry calibration, a 50-basis-point decline in annualised loan returns reduces wholesale funding by about 1.6% on impact, compared with just over 0.5% with a representative depository bank, highlighting the role of heterogeneity in credit supply responses.

*University College London. Email: allen.shen.16@ucl.ac.uk.

1 Introduction

A bank can finance a mortgage it never originates. U.S. banks supply warehouse credit to nonbank mortgage lenders, including firms they compete with for borrowers (Jiang, 2019). When a bank takes losses and has to shrink its balance sheet, does it make fewer loans itself or withdraw funding from those lenders? I show that the answer depends on the size of the bank that bears the loss. In the model, small banks cut both forms of lending. Large banks hold liquid claims on other intermediaries above liquidity regulation requirements, giving them room to cut wholesale funding first. In a calibrated loan-loss experiment, accounting for this difference nearly triples the initial fall in wholesale funding relative to a representative-bank benchmark.

The motivating fact is a pattern in U.S. bank portfolios. Figure 1 pools quarterly Call Reports from 2013Q1 to 2021Q4 and plots the share of assets invested in mortgage-backed and other asset-backed securities against bank size. The share stays close to 6% over much of the size distribution, then rises steeply, reaching nearly 15% in the largest size bin. These holdings provide a proxy for banks' investment in claims that can finance credit intermediation and provide liquidity. The open question is whether the banks with the largest holdings also draw on them most when they lose capital.

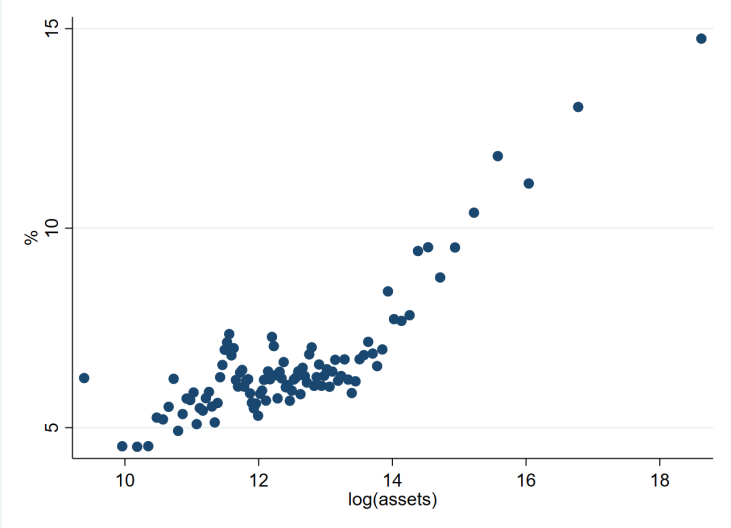


Figure 1. Securitised claims as a share of assets, by bank size

Notes. Each point is a size bin in pooled quarterly U.S. bank Call Reports, 2013Q1–2021Q4. The vertical axis measures holdings of mortgage-backed securities (including commercial mortgage-backed securities) and other asset-backed securities as a percentage of total assets. The horizontal axis is the log of total assets.

In the model, the same claim appears on two balance sheets. It funds loans at a wholesale intermediary that cannot issue insured deposits. The depository that holds the claim earns a return and can use it to meet a liquidity requirement. I refer to the claim as a private safe asset in this sense. The bank can instead make loans directly. The marginal return on

those loans falls as the bank exhausts its most profitable lending opportunities. At a given wholesale rate, there is therefore a loan-book size at which making one more loan is just as profitable as supplying wholesale funding.

That indifference point separates two kinds of banks. A small bank would like to make more direct loans, but must hold a fraction of its balance sheet in liquid claims. When it loses a dollar of balance-sheet capacity, direct loans and wholesale claims both fall in proportion. A sufficiently large bank has already made every direct loan that yields more than the wholesale rate. It holds wholesale claims above the liquidity minimum, so at unchanged returns the whole dollar comes out of wholesale funding. Size matters because it determines the use of the *marginal* dollar. A loss concentrated at banks above the threshold withdraws more funding from nonbank lenders per dollar of bank capital lost.

I quantify this mechanism by comparing an economy with a representative depository bank to one with large and small depository banks. The two economies have similar aggregate bank net worth. In steady state, the wholesale sector accounts for 59.3% of direct lending to final borrowers in the heterogeneous economy, compared with 51.3% in the representative-bank economy. It also experiences a sharper funding contraction after losses. A 50-basis-point decline in the annualised loan return reduces wholesale funding on impact by about 1.6% with heterogeneous banks and just over 0.5% with a representative bank. Yet the wholesale rate falls *less* in the heterogeneous economy. The wholesale intermediary loses net worth and borrows less in both economies. When large depositories also withdraw funding, the rate need not fall as far to clear the market.

Related literature. The closest theoretical benchmark is Gertler et al. (2016), who study wholesale funding and runs in a banking system with retail and wholesale intermediaries. I ask how a loan loss propagates through that funding market when the depositories supplying funds differ in size. The mechanism operates even without a run. It complements Begenau and Landvoigt (2021), who study how capital regulation reallocates activity between commercial and shadow banks, and Coimbra and Rey (2017), who link the distribution of intermediary risk-taking to financial cycles. Here, banks with the same lending technology respond differently because they hold different amounts of wholesale claims above their liquidity minimum.

The paper also connects empirical work on bank–nonbank funding to the literature on private safe assets. Jiang (2019) documents banks' financing of competing mortgage lenders. **?** show more broadly that banks and nonbanks fund one another, with nonbanks particularly reliant on banks. A separate literature documents the liquidity services of shadow bank debt and the convenience yield on agency mortgage-backed securities (Sunderam, 2015; He and Song, 2022). I connect the holder's demand for liquidity to the issuer's funding decision and derive how a loss at one intermediary affects lending at the other.

2 Model

Time is discrete. A unit mass of depository banks, indexed by j, raises insured deposits at an exogenous gross rate R^d. Each bank chooses between lending directly to final borrowers and acquiring claims on a representative wholesale bank. The wholesale bank also lends to final borrowers, but relies on the depositories for external finance. Let R^b denote the gross return on a wholesale claim. The wholesale funding market determines R^b.

A banker of type $i \in \{D, W\}$ survives each period with probability σ^i. A banker who exits consumes the bank's net worth and is replaced by an entrant. Depository entrants can differ in their endowments, ω^j, but face the same deposit rate, loan technology and regulatory parameters. The endowment of an entering wholesale banker is ω^W. Bankers discount future payoffs by a factor β. For either type, the value of operating a bank with net worth n_t is

$$V_t(n_t) = \max \ \beta \, \mathbb{E}_t \left[(1 - \sigma^i) n_{t+1} + \sigma^i V_{t+1}(n_{t+1}) \right]. \tag{1}$$

The feasible portfolios and the law of motion for net worth differ across the two types.

2.1 Depository banks

At date t, bank j chooses its direct loan book k^j_{t+1}, wholesale claims b^j_{t+1} and deposits d^j_{t+1}. The subscript on a position denotes its payoff date. Total assets A^j_t and next period's net worth satisfy

$$A^j_t \equiv k^j_{t+1} + b^j_{t+1} = n^j_t + d^j_{t+1}, \tag{2}$$

$$n^j_{t+1} = F^D_{t+1}(k^j_{t+1}) + R^b_{t+1} b^j_{t+1} - R^d d^j_{t+1}. \tag{3}$$

The total payoff $F^D(k)$ is increasing over the relevant range and strictly concave. A bank makes its most profitable loans first. Expanding its loan book lowers the return on the next loan, whereas it can supply wholesale funds at the common market rate.

Two constraints determine how the bank can use its net worth.

$$n^j_t \geq \theta A^j_t, \tag{4}$$

$$b^j_{t+1} \geq \rho d^j_{t+1}. \tag{5}$$

The first requires the bank to finance a fraction θ of assets with equity. The second requires a stock of liquid claims against deposits. It captures the liquidity service of the model's private safe asset. The Basel liquidity coverage ratio motivates this constraint by relating eligible high-quality liquid assets to stressed net cash outflows over 30 days (Bank for International Settlements, 2013). The model compresses asset eligibility, haircuts and outflow rates into one parameter, ρ. The floor can also capture an internal liquidity

policy. A warehouse credit commitment illustrates the funding link, while a liquid security illustrates a claim that can meet the holder's liquidity needs.

The capital constraint binds in the calibrated equilibria. At steady-state rates, $R^b > R^d$ is sufficient. A bank with spare capital can raise a dollar of deposits and invest it in a wholesale claim at a positive spread while satisfying the liquidity requirement ($\rho < 1$). Thus $A = n/\theta$ and deposits equal $(1-\theta)A$. Writing $\ell \equiv \rho(1-\theta)$, the liquidity constraint requires $b \geq \ell A$, or equivalently $k \leq (1-\ell)A$.

Portfolio choice and the size threshold. Consider a steady state and hold returns fixed. Let $k^\star$ solve $(F^D)'(k^\star) = R^b$, with $k^\star > 0$ as in the calibration. The bank prefers direct loans up to $k^\star$ and wholesale claims thereafter. Its optimal portfolio is

$$k(A) = \min\{(1-\ell)A,\, k^\star\}, \qquad b(A) = A - k(A) = \max\{\ell A,\, A - k^\star\}. \tag{6}$$

The threshold in assets is $A^\star = k^\star/(1-\ell)$, or $n^\star = \theta A^\star$ in net worth. Below it, the liquidity constraint binds and the wholesale share of assets equals ℓ. Above it, the bank holds more than the required minimum and the share equals $1 - k^\star/A$, which rises with size. This is the sense in which the model reproduces the shape in Figure 1.

Equation (6) also gives the response to a loss. For a dollar of assets lost at unchanged returns, wholesale funding falls by ℓ dollars below the threshold and by one dollar above it, provided the bank stays in the same regime. Moreover, $b(A)$ is convex. At a common wholesale rate, banks on opposite sides of the threshold supply more wholesale funding in total than a representative bank with their mean assets and the same loan technology. This is a statement about the supply schedule. In equilibrium the wholesale rate moves, changing $k^\star$ and the quantity borrowers demand.

2.2 The wholesale bank

The wholesale bank combines its own net worth n_t^W with borrowing B_{t+1} to finance loans k_{t+1}^W. Its specialised lending technology pays a common gross return R_{t+1}^k on each loan. The balance sheet and net worth of a continuing bank satisfy

$$k_{t+1}^W = n_t^W + B_{t+1}, \tag{7}$$

$$n_{t+1}^W = R_{t+1}^k k_{t+1}^W - R_{t+1}^b B_{t+1}. \tag{8}$$

Following Gertler and Karadi (2011) and Gertler et al. (2016), limited commitment restricts external finance. After raising funds, the banker can divert a fraction γ_1 of assets net of the recoverable portion $(1-\gamma_2)B_{t+1}$ of its debt. Creditors lend only if continuation is at

least as valuable as diversion,

$$V_t^W(n_t^W) \;\geq\; \gamma_1\big[k_{t+1}^W - (1-\gamma_2)B_{t+1}\big]. \tag{9}$$

The linear loan technology and proportional diversion payoff make the wholesale bank's value proportional to net worth at given prices. With a binding incentive constraint, its leverage is

$$\phi_t \equiv \frac{k_{t+1}^W}{n_t^W} = \frac{\mathbb{E}_t[\Omega_{t+1}R_{t+1}^b] - \gamma_1(1-\gamma_2)}{\gamma_1\gamma_2 - \mathbb{E}_t[\Omega_{t+1}(R_{t+1}^k - R_{t+1}^b)]}, \tag{10}$$

where $\Omega_{t+1} \equiv \beta[1-\sigma^W + \sigma^W V_{t+1}^W(n_{t+1}^W)/n_{t+1}^W]$. Wholesale borrowing is therefore $B_{t+1} = (\phi_t - 1)n_t^W$. A fall in net worth reduces the scale on which the intermediary can borrow for a given leverage ratio.

2.3 Aggregation and equilibrium

Sectoral net worth includes surviving bankers' retained earnings and entrants' endowments. If h_t^i denotes the next-period net worth of a surviving bank under its optimal policies and $N_t^D = \int n_t^j\, dj$, the sector totals obey

$$\begin{aligned} N_{t+1}^D &= \sigma^D \int h_t^D(n_t^j)\, dj + (1-\sigma^D)\int \omega^j\, dj, \\ N_{t+1}^W &= \sigma^W h_t^W(N_t^W) + (1-\sigma^W)\omega^W. \end{aligned}$$

Given loan returns and the deposit rate, banks choose feasible portfolios to maximise (1). The equilibrium wholesale rate equates funding supplied by depositories with borrowing by the wholesale bank,

$$\int b_{t+1}^j\, dj = B_{t+1}. \tag{11}$$

Total credit to final borrowers is $K_{t+1} = \int k_{t+1}^j\, dj + k_{t+1}^W$. Wholesale claims appear once on each side of the financial sector's consolidated balance sheet. Counting them as additional final loans would double-count the credit they finance.

3 Calibration

One period is a quarter. I compare two economies that share the loan technology, deposit rate, regulatory constraints and wholesale banking sector. The benchmark represents depositories with a single balance sheet. The second economy distinguishes large and small depositories. This comparison asks what the distribution of bank net worth adds to a description based on sector totals.

The total loan payoff for intermediary type i is

$$F_t^i(k) = R_t^k k - \frac{\alpha_1^i}{2} k^{\alpha_2}.$$

I set $\alpha_2 = 2$, $\alpha_1^D = 0.0099$ and $\alpha_1^W = 0$. At fixed returns the depository's interior loan book is therefore $k^\star = (R^k - R^b)/\alpha_1^D$. Both rates in this expression are quarterly gross rates. Table 1 reports the steady-state loan and deposit rates in annualised terms.

Table 1. Bank parameters and steady-state returns

Parameter	Interpretation	Value
β	Quarterly discount factor	0.99
$\sigma^D,\ \sigma^W$	Survival probabilities, depository and wholesale	0.88, 0.62
$\omega^D,\ \omega^W$	Entrant endowments, benchmark depository and wholesale	0.0088, 0.0169
$\omega^B,\ \omega^S$	Depository entrant endowments, large and small	0.0242, 0.0040
$\gamma_1,\ \gamma_2$	Diversion share and unrecoverable share of wholesale debt	0.06, 1
$\alpha_1^D,\ \alpha_1^W$	Loan cost coefficients, depository and wholesale	0.0099, 0
α_2	Loan cost exponent	2
θ	Minimum equity as a fraction of assets	0.10
ρ	Minimum wholesale claims as a fraction of deposits	0.5556
$R^k,\ R^d$	Annualised gross loan and deposit returns	1.064, 1.038

The endowments target a steady-state ratio of depository to wholesale net worth of roughly four to one. The large type represents the largest 5% of U.S. depositories, which account for nearly 80% of depository net worth in the calibration target. The small type represents the remainder. Endowments differ to capture this concentration. The wholesale diversion parameters discipline leverage, with $\gamma_2 = 1$ for simplicity.

The liquidity requirement is particularly important for interpreting the exercise. Setting $\theta = 0.10$ and $\rho = 0.5556$ implies that a constrained bank holds approximately half its assets in the model's wholesale claim. That claim represents the bank's entire liquid portfolio in a two-asset model. The securities plotted in Figure 1 cover only part of such a portfolio. I use Figure 1 to discipline the shape of the portfolio rule. The assumption that all liquid holdings finance the wholesale sector also matters for the magnitude of the simulated funding withdrawal.

For the dynamic experiment I apply the same temporary decline in R_t^k to both economies, starting from their respective steady states. The initial decline is 50 basis points at an annual rate, and the subsequent return path is shown in Figure 2. The shock lowers the payoff on outstanding direct loans in both sectors. The deposit rate and the parameters governing capital and liquidity requirements stay fixed.

4 Results

4.1 Who supplies credit in steady state?

The two economies have similar aggregate bank net worth but allocate credit differently (Table 2). The representative depository holds the minimum liquid share, 50.0%. In the heterogeneous economy the small bank also sits at the minimum, while the large bank holds claims above it. Wholesale claims consequently account for 58.2% of aggregate depository assets. With this additional supply, the wholesale bank pays 4.37% annually, 43 basis points below the representative-bank benchmark. Its leverage increases from 21.00 to 22.93, and its share of lending to final borrowers rises by 8.0 percentage points, from 51.3% to 59.3%.

Table 2. Steady states under the two depository banking structures

	Representative	Heterogeneous
Depository net worth, N^D	0.08	0.08
Wholesale net worth, N^W	0.02	0.02
Loan return at zero scale, annualised net rate	6.40%	6.40%
Wholesale funding rate, annualised net rate	4.80%	4.37%
Wholesale share of loans to final borrowers	51.27%	59.30%
Depository share of loans to final borrowers	48.73%	40.70%
Wholesale claims as a share of depository assets	50.00%	58.17%
Wholesale leverage, ϕ	21.00	22.93

4.2 How does wholesale funding respond to a loan loss?

Figure 2 reports deviations from each economy's own steady state. The same decline in loan returns reduces net worth at both types of intermediary. On impact, wholesale funding falls about 1.6% in the heterogeneous economy and just over 0.5% in the benchmark. The difference is sharpest at the start. Funding reaches a trough of approximately 1.9% below steady state in the heterogeneous economy, compared with about 1.4% in the benchmark. The representative economy reaches its trough later and recovers more slowly.

The portfolio responses explain the difference. When the representative bank loses net worth, its binding capital constraint shrinks its balance sheet. Its binding liquidity constraint keeps the portfolio shares fixed, so direct loans and wholesale claims contract together. The large bank in the heterogeneous economy starts above the liquidity minimum. At given returns its direct loan book is determined by $k^\star$, so wholesale claims absorb the loss of balance-sheet capacity. Equilibrium returns move as well, but the portfolio implication is visible in the bottom row of Figure 2. The wholesale share of depository assets falls by roughly 0.7 percentage points on impact with heterogeneity and stays flat in

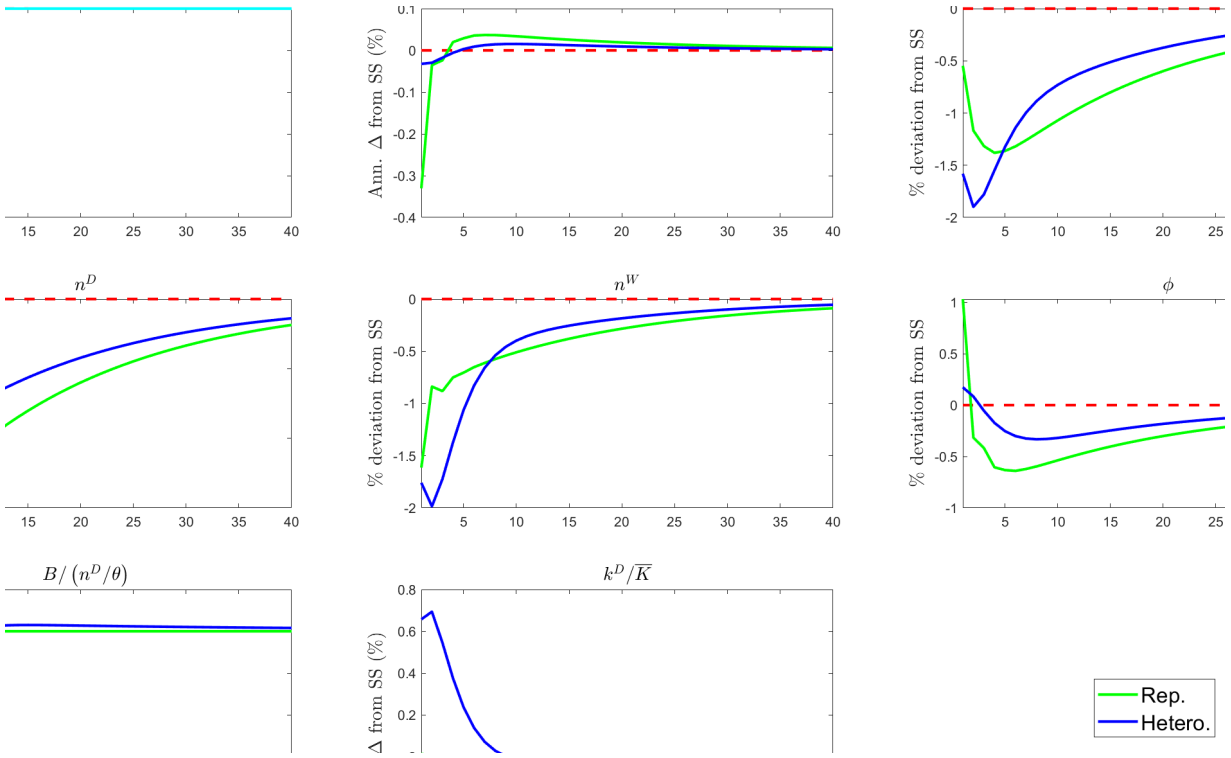


Figure 2. Responses to a 50-basis-point fall in the annualised loan return

Notes. Quarters since the shock run along the horizontal axes, beginning with the impact quarter. The blue line denotes the heterogeneous economy, the green line the representative economy and the red dashed line zero. The common loan return is in cyan. Rate changes are annualised percentage-point differences. Funding, net worth and leverage are percentage deviations. Portfolio and intermediation shares are percentage-point differences.

the benchmark. Depositories' share of final lending rises as the wholesale bank's lending contracts disproportionately.

The price response runs in the opposite direction to the quantity comparison. The annualised wholesale rate falls on impact by roughly 3 basis points with heterogeneous banks and about 30 basis points in the representative economy. Both sides of the wholesale market help explain the pattern. The wholesale bank loses net worth and demands less external finance, even though its leverage initially rises. This puts downward pressure on the funding rate. Withdrawal of funds by the large depository puts upward pressure on the same rate. The stronger withdrawal of supply in the heterogeneous economy partly offsets the effect of weaker demand on the funding rate. A modest rate movement can therefore accompany a substantial withdrawal of credit from nonbank lenders.

5 Conclusion

Which banks bear a loss matters for the lenders they finance. A depository at its liquidity minimum cuts direct and wholesale lending together. A bank with a large enough loan book holds liquid claims above that minimum. At given returns it can shrink its balance sheet entirely by reducing funds supplied to other lenders. This distinction changes the interpretation of a wholesale funding rate. In the calibrated economy with heterogeneous banks, nonbank funding contracts more on impact even as its price moves less.

The mechanism also suggests a test that goes beyond the cross-sectional pattern in the Call Reports. Consider two banks that fund the same nonbank lender and experience comparable losses on loans made before the shock. Suppose one bank entered the shock at its liquidity minimum and the other held claims in excess of it. For a small contraction at common returns, the model predicts a larger cut in wholesale funding per dollar of assets lost at the second bank. Comparing funding changes within the same borrower would help separate the withdrawal of credit supply from a decline in the borrower's demand. Distinguishing net purchases or new advances from valuation changes would then show whether banks actively reallocate their portfolios in the way the model predicts.

References

Bank for International Settlements. Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools. Technical report, Basel Committee on Banking Supervision, 2013. URL `https://www.bis.org/publ/bcbs238.pdf`.

J Begenau and T Landvoigt. Financial Regulation in a Quantitative Model of the Modern Banking System. *The Review of Economic Studies*, 12 2021. ISSN 0034-6527. doi: 10.1093/restud/rdab088.

Nuno Coimbra and Hélène Rey. Financial Cycles with Heterogeneous Intermediaries. 2017.

M. Gertler, N. Kiyotaki, and A. Prestipino. Wholesale Banking and Bank Runs in Macroeconomic Modeling of Financial Crises. pages 1345–1425. 2016. doi: 10.1016/bs.hesmac.2016.03.009.

Mark Gertler and Peter Karadi. A model of unconventional monetary policy. *Journal of Monetary Economics*, 2011. ISSN 03043932. doi: 10.1016/j.jmoneco.2010.10.004.

Zhiguo He and Zhaogang Song. Agency MBS as Safe Assets. 4 2022. URL `https://www.nber.org/papers/w29899`.

Erica Xuewei Jiang. Financing Competitors: Shadow Banks' Funding and Mortgage Market Competition. 12 2019.

Adi Sunderam. Money Creation and the Shadow Banking System. *The Review of Financial Studies*, 28(4):939–977, 4 2015. ISSN 0893-9454. doi: 10.1093/RFS/HHU083. URL `https://academic.oup.com/rfs/article/28/4/939/1928049`.