the review of income and wealth

Review of Income and Wealth
Series 0, Number 0, 2024
DOI: 10.1111/roiw.12689

CASH HOLDINGS AND HEALTH SHOCKS

BY LIN TIAN

Aston University

ALESSANDRA GUARIGLIA* and NICHOLAS HORSEWOOD

University of Birmingham

We are the first to investigate how health shocks relate to cash holdings. Using three waves of the China Health and Retirement Longitudinal Study over the period 2013–2018, we document that, for middle-aged and elderly people living in rural China, the onset of an acute health condition is associated with a 3.0 percentage point higher probability of holding only cash as a safe asset, and a 2.3 percentage point higher proportion of safe assets held in the form of cash. These results are robust to using different samples and estimation methods. We also find that ex-post reimbursement of medical expenses and lack of bank accessibility may drive the association between health shocks and cash holdings.

JEL Codes: D14, I13, I14

Keywords: banks, cash holdings, China, health insurance, health shocks

1. INTRODUCTION

In recent years, unprecedented financial innovation and credit expansion have provided households with access to a broad array of financial services and products. However, empirical studies find that despite the high returns associated with risky assets, a large fraction of people choose not to hold them in their portfolios (Calvet et al., 2009; Ivković et al., 2008).

An extensive literature explores various determinants of portfolio choice (Calvet & Sodini, 2014; Christelis et al., 2010; Heaton & Lucas, 2000). Health status and health shocks have been considered among such determinants (Berkowitz & Qiu, 2006; Coile & Milligan, 2009; Døskeland & Kvaerner, 2022; Goldman & Maestas, 2013; Rosen & Wu, 2004). A poor health status in general and health shocks in particular can cause significant out-of-pocket medical expenditures and

Note: We thank the Editor, Conchita D'Ambrosio, and two anonymous referees for excellent comments and a very professional review process. We are also grateful to participants at various conferences and seminars for insightful discussions and valuable feedback on earlier versions of this paper.

*Correspondence to: Alessandra Guariglia, Department of Economics, Birmingham Business School, University of Birmingham, University House, Edgbaston, Birmingham B15 2TY, UK (a.guariglia@bham.ac.uk).

income loss in the short run and impair wealth accumulation in the long run. In line with this argument, Dobkin et al. (2018) document that, as a consequence of hospitalization (which can be seen as the result of a sharp health shock), non-elderly adults in the US experience an increase in out-of-pocket medical spending, together with a considerable reduction in earnings, access to credit, and consumer borrowing. As people experiencing health shocks face significant uncertainty not knowing how much they will need to pay for their treatment and when, and how much of their income will be lost as a result of their illness, they have been found to reduce their exposure to other financial risk, leading to safer portfolios (Berkowitz & Qiu, 2006; Rosen & Wu, 2004). Additionally, a health shock could affect risk preferences, which could in turn lead to a safer portfolio (Decker & Schmitz, 2016). The onset of new health conditions may also affect portfolio choice through a revision of survival expectations (Døskeland & Kvaerner, 2022) or through a change in time preferences and planning horizons (Hong & Hanna, 2014).

Whilst most of this literature has looked at portfolio choice focusing on risky assets holdings, to the best of our knowledge, no paper has studied cash holdings. We fill this gap by analyzing, for the first time, how health shocks relate to cash holdings. To this end, we focus on the middle-aged and elderly Chinese rural population using data drawn from the China Health and Retirement Study (CHARLS) over the period 2013–2018.

We believe rural China represents an ideal laboratory to study the effect of health shocks on cash holdings for the following reasons. First, the Chinese financial system is still in the initial stages of development. As a result, by 2017, China had the world's largest unbanked population with 225 million of adults not having a bank account. This problem was particularly severe in rural China, where 21 percent of adults did not have a bank account, meaning that they only held cash in their portfolios (Demirgüç-Kunt et al., 2018). Even if they do hold a bank account, many Chinese rural residents still prefer to hold and use cash, instead of using cards and/or mobile payments. Furthermore, according to Sui and Niu (2018), less than one percent of the rural Chinese population holds risky assets (i.e., stocks, bonds, and mutual funds).

Second, the Chinese population, and, in particular, the older population, is not healthy, and, as such, particularly vulnerable to adverse health conditions: in 2019, around 75 percent of older people (aged 60 and above) in China were found to suffer from noncommunicable disease such as cardiovascular disease, diabetes, and hypertension (WHO, 2021).

Third, although the health insurance system has been developed substantially in the past decades, a considerable fraction of Chinese people still have to pay high out-of-pocket (OOP) medical expenses (Wang et al., 2018). The problem is exacerbated in rural areas (Lei & Lin, 2009; Wang et al., 2016; Zhang et al., 2017). As a result, health shocks may have a significant impact on Chinese people's financial situation, which, in turn, can affect their portfolio choice.

Our paper also contributes to the literature on the consequences of adverse health shocks. Many papers within this literature have looked at how health shocks affect labor supply (García-Gómez, 2011; Gupta et al., 2015; Jones et al., 2020; Riekhoff & Vaalavuo, 2021). Others have analyzed how they relate to unhealthy behaviors such as smoking and/or obesity (Clark & Etilé, 2002; Sundmacher, 2012), charitable giving (Black et al., 2021), healthcare use (Fiebig et al., 2021), and household spending and income (Cheng et al., 2019). Finally, as mentioned above, a few papers have linked health shocks to people's willingness to hold risky assets (Berkowitz & Qiu, 2006; Døskeland & Kvaerner, 2022; Rosen & Wu, 2004). We advance this literature by looking, for the first time, at how health shocks relate to cash holdings. We also investigate possible drivers of the association between health shocks and cash holdings, focusing on the role of the type of reimbursement of healthcare expenses that patients receive from their insurance scheme and the role of bank accessibility.

Understanding the determinants of cash holdings in rural China can help policymakers identify possible tools to be used to reduce excessive reliance on cash. This is particularly important as a high reliance on savings in the form of cash in an economy has the potential for adverse effects on economic development (Levine & Zervos, 1998; Stix, 2013): money held in the form of cash is in fact not deposited in financial institutions and lent out to borrowers for investment projects. Moreover, too much cash in an economy has been associated with crime, tax evasion, and activities in the underground economy (Lahiri, 2020). Finally, from the perspective of the household, cash is hard to manage, prevents wealth accumulation in the long run (Demirgüç-Kunt et al., 2018), can be easily lost or stolen, and does not have any record-keeping capability in which consumers can trace back their payments (Shy, 2023). More specifically, understanding what drives the association between health shocks and cash holdings can provide policymakers with an indication of possible ways to encourage people facing health shocks to hold assets in more productive ways than cash.

The remainder of the paper is structured as follows. Section 2 describes the institutional setting, focusing on the public health insurance system and financial inclusion in China. Section 3 presents our testable hypothesis. Section 4 describes our data and reports some descriptive statistics. Section 5 illustrates our baseline specification and estimation methodology. Section 6 presents our main results and Section 7, a series of tests aimed at ascertaining that the relationship between health shocks and cash holdings is causal. Section 8 discusses the role of reimbursement method and bank accessibility as possible drivers of the relationship. Section 9 concludes.

2. Institutional Setting

2.1. *Public Health Insurance System in China*

After decades of endeavor, China achieved universal coverage by providing health insurance to over 95 percent of the population by 2011. The health insurance system in China comprises three major health insurance programs, namely, the Urban Employee Basic Medical Insurance (UEBMI), the Urban Resident Basic Medical Insurance (URBMI), and the New Rural Cooperative Medical Scheme (NRCMS). The first two programs aim at urban residents with contributions on either a compulsory or voluntary basis, while the NRCMS is designed for rural residents on a voluntary basis. There are significant differences across the three programs in the share of the burden of inpatient care to be covered by the insured: Ma and Cen (2017) document that the percentage of the health care fee paid by the individual (or household) is 56.0 percent for rural residents, whereas it ranges between 31.8 and 32.8 percent for urban residents.

Since we focus on rural residents, the relevant insurance scheme is the NRCMS, which was launched in 2003 and implemented in 2008. The scheme is administrated by local governments and is based on a voluntary participation system and individual contributions. Although it is a voluntary scheme, due to the modest program premiums, relatively generous government subsidies, and strong government mobilization ability, it is convenient for the vast majority of rural residents to participate (Zhong, 2011). According to the National Health and Family Planning Commission, P.R. China (2014), by 2014, the scheme covered more than 800 million rural Chinese, accounting for about 98.7 percent of the total rural population.

Yet, the NRCMS mainly aims at relieving and hedging inpatients from catastrophic disbursement of hospital expenses, with outpatient expenses only (partially) covered in a limited number of counties. Furthermore, the scheme only covers 800–1200 drugs, meaning that if other drugs are prescribed, patients need to pay for them out of their own pocket (Wang et al., 2016). Finally, according to our dataset, around 30 percent of NRCMS beneficiaries get immediate reimbursement of their medical expenditures (i.e., they only pay at the coinsurance rate for each treatment exceeding the deductible), while the remaining 70 percent have to pay the full amount of the health care costs up-front and get reimbursement later (Lai et al., 2018; Zhong, 2011). The complexity of cost sharing arrangements may cause confusion among the insured about what can and cannot be reimbursed, how much will be reimbursed, and when the reimbursement will take place (Zhong, 2011).

2.2. *Financial Inclusion in Rural China*

The Chinese government started to commit to financial inclusion in the late 20th century, with intensified effort since 2005, paying particular attention to supporting small- and middle-sized enterprises and rural residents (Peng et al., 2014). Yet, by 2017, in rural China, around 21 percent of the population aged 15 and over did not have a bank account (Demirgüç-Kunt et al., 2018).

According to the 2013–2018 waves of the CHARLS, less than one percent of rural respondents aged 45 and over hold bonds, stocks, and/or mutual funds. A number of papers have investigated reasons why Chinese households hold a small amount of risky assets. One reason is the low financial development characterizing China (Allen et al., 2017). A second reason is that the Shanghai Stock Exchange Index has had average annual real returns of zero from 2001 to 2016 (Glaeser et al., 2017). A third reason is the low financial literacy characterizing Chinese households (Feng et al., 2019; Zou & Deng, 2019). The problem of low risky asset holding is exacerbated among older adults living in rural areas, which are the focus of our analysis. This can be explained bearing in mind that older Chinese adults show a particularly low education attainment, as well as low levels of financial literacy (Li et al., 2020). Similarly, people living in rural areas are characterized by low education attainment and low income and wealth, which are additional factors explaining their low risky asset holding (Sicular et al., 2007; Sui & Niu, 2018).

The 2013–2018 waves of the CHARLS also show that among respondents who only hold bank deposits and/or cash, 67.6 percent hold only cash, implying that they do not have or use a bank account. As a result, cash retains a dominant position, especially for the older and less "tech-savvy" population living in rural areas.

3. Testable Hypothesis

A health shock may lead to significant income losses and to an increase in non-medical spending resulting from changes in life circumstances due to poor health (De Nardi et al., 2010; Wu, 2003). As a result of the uncertainties surrounding both income losses and non-medical expenditures, people suffering from health shocks have been found to reduce their exposure to other financial risks, and, hence, to hold safer portfolios (Berkowitz & Qiu, 2006; Coile & Milligan, 2009; De Nardi et al., 2010; Døskeland & Kvaerner, 2022). In the Chinese context, these uncertainties are magnified as people do not know how much they will need to pay for their treatment upfront and whether and when they will receive a reimbursement (Zhong, 2011). As a consequence, Chinese people prefer to be relatively liquid.

Furthermore, in the context of rural China, absent of a well-developed financial market and adequate financial education, residents are unfamiliar with financial services and trust in banks is limited (Fungáčová & Weill, 2018). Consequently, people wishing to hold a safer portfolio will hold more cash. Especially for older and less tech-savvy people, cash is in fact considered as the safest of safe assets.

In line with these considerations, we hypothesize that, in the context of rural China, people suffering from a health shock will be less willing to take risks with their financial wealth. They will therefore hold a higher proportion of their safe assets in the form of cash (rather than putting money in their bank account) and will be more likely to hold only cash as a safe asset.

4. Data

4.1. *The China Health and Retirement Longitudinal Study*

We use the China Health and Retirement Longitudinal Study (CHARLS), which is a nationally representative panel survey of people aged 45 years and older and their spouses from around 450 communities across 28 Chinese provinces. Wave 1 (baseline survey) was conducted in 2011, and waves 2, 3, and 4, respectively in 2013, 2015, and 2018.

The dataset provides a wide range of information, including socio-demographics, household consumption and income. It also provides detailed information on health and household financial assets. The 2011 wave additionally provides detailed information on community infrastructures. As it does not provide information on health shocks, we do not use this wave in our main analysis.[1]

[1] However, we use the 2011 wave of the survey when estimating models that make use of lagged variables, namely the models based on Propensity Score Matching and entropy balancing and the Generalized Methods of Moments models. These are described in Sections 7.2 and 7.3.

After retrieving the data from different modules and merging them based on household and individual identifiers, we obtain a preliminary three-period panel that contains 56,323 observations for respondents aged 45 and over and their spouses. Since our study is centered on respondents who live in rural areas, we exclude all respondents living in urban areas. This leaves us with 33,468 observations. We further exclude 524 respondents with negative net household income, and respondents with missing observations for the variables included in our model.[2] We end up with a three-period unbalanced panel made up of 18,198 observations for 9030 individuals. Among these, 2895 individuals participate in all three waves of the survey. Furthermore, we winsorise the continuous control variables used in our models (i.e., assets and income) at the 1 percent level of both tails to mitigate the influence of outliers. These same variables are converted into real terms using the province-level Consumer Price Index (CPI) in rural areas calculated by the National Bureau of Statistics of China, using 2011 as the base year.

4.2. *Key Variables*

We consider two main outcome variables. The first, *prcash*, is a dummy variable which equals 1 if the respondent holds only cash as a safe asset, and 0 otherwise. The second, *cashratio*, is the proportion of safe assets held by the respondent in the form of cash and is measured as the ratio of cash over the sum of cash and bank deposits. We focus on safe assets because, as noted above, these are the most liquid and accessible among the financial assets available to respondents. Moreover, according to our dataset, more than 99 percent of rural respondents keep all their financial assets in the form of cash or bank deposits.

Our key explanatory variable is an objective health shock. Specifically, we construct a dummy variable, *Acute*, which equals to 1 if the respondent suffered the onset of an acute illness such as cancer, heart disease, stroke, diabetes, and chronic lung disease in the past 2 years, and 0 otherwise (Cheng et al., 2019; Clark & Etilé, 2002; Døskeland & Kvaerner, 2022; Gupta et al., 2015; Jones et al., 2020; Riekhoff & Vaalavuo, 2021; Smith, 2005).

Although our health shock is self-reported, and as such prone to justification bias (Anderson & Burkhauser, 1985), individuals have been found to be less likely to misreport the presence or new diagnosis of specific conditions (Gupta et al., 2015). Furthermore, even though health conditions may be correlated with individuals' lifestyle and other socio-economic status, the occurrence and timing of such conditions is likely to be unpredictable. However, in Section 7.3.2, we present results in which the *Acute* dummy is instrumented.

4.3. *Descriptive Statistics*

Appendix Table S1 presents descriptive statistics for the main variables used in the paper, which are defined in Table 1. Overall, 67.6 percent of our sample hold

[2] A relatively large number of observations show missing cash holdings. Our findings were robust to imputing missing cash holdings using the fitted values from a regression model, as well as to using the multiple imputation by chained equations (MICE) approach as in Shoji et al. (2022). These results are not reported for brevity but are available upon request.

only cash as a safe asset. Furthermore, on average, the people in our sample hold 73.0 percent of their safe assets in the form of cash. As for health shocks, 6.3 percent of respondents were diagnosed with new acute conditions in the last 2 years.

Table 2 shows that the mean values of *prcash* (Panel A) and *cashratio* (Panel B) are higher for respondents who suffered from a health shock, compared to those who did not (column 1 and 6). The differences in means are statistically significant for both outcome variables. In the following sections, we provide a formal econometric analysis on the links between cash holdings and health shocks.

Appendix Table S1 also shows that respondents in our sample have an average age of 61, with half of the respondents being male and 88.5 percent of the respondents being married. About one in 10 of the households in our sample have more than two adult members. Just under 27 percent of the respondents are illiterate, and less than 8 percent hold qualifications of high school or above. Around 16 percent of the respondents are employed, 61.6 percent engage in agricultural work, are self-employed, or do unpaid housework, while 22.1 percent do not work at all. The average score of cognitive ability is 11 out of 21. Just under 88 percent of the respondents own at least one house. Though computers are not prevalent in the rural areas, 16.9 percent of the respondents own a computer in their home and just under nine in 10 have mobile phones. Household income is right skewed, with an average at about 22,400 CNY and a median of about 9634 CNY. The amount of household financial assets is also widely dispersed: the average amount is 17,029 CNY, whilst the median is only 1846 CNY. Furthermore, 9.3 percent of the respondents live in communities with at least one bank branch, and 32.1 percent of the respondents are eligible for immediate reimbursements through the health insurance schemes they are enrolled in.

5. Baseline Specification and Estimation Methodology

5.1. *Baseline Model*

In order formally to test our hypothesis, we initially estimate the following model, where i indexes individuals; p, the provinces where they reside; and t, time:

$$Y_{i,p,t} = \beta_1 Acute_{i,p,t} + \gamma' X_{i,p,t} + \delta' Z_{i,p} + \varphi_p + \varphi_t + \varphi_i + \varepsilon_{i,p,t} \ (i = 1, \dots, \mathrm{N};\ t = 1, 2, 3;\ p = 1, \dots, 24) \quad (1)$$

$Y_{i,p,t}$ represents in turn one of the two outcome variables we consider, namely *prcash* and *cashratio*. $Acute_{i,p,t}$ refers to the occurrence of an acute health shock to individual i in the past 2 years. $X_{i,p,t}$ denotes a set of time-varying control variables including the respondent's age, age squared, marital status, educational attainment, employment status, cognitive ability, household size; dummy variables indicating home, computer, and mobile phone ownership; household income, household wealth, and health insurance reimbursement method. $Z_{i,p}$ is a vector of time-invariant control variables including gender and bank accessibility.

Our selection of control variables in Equation (1) is inspired by the literature on household portfolio choice. A number of studies find that older individuals are less likely to invest in risky assets, since they are more risk-averse than younger generations (Calvet & Sodini, 2014; Cardak & Wilkins, 2009). Yet, Ameriks and Zeldes (2000) find a non-linear relationship between age and equity holdings. Gender differences exist in financial portfolio choice, primarily through disparity in risk attitude and financial knowledge: men are expected to be less risk-averse and more financially literate than women and hence hold riskier financial portfolios (Charness & Gneezy, 2012; Halko et al., 2012). In light of this literature, we expect males and younger respondents to rely less on cash. In line with Ameriks and Zeldes (2000), we also allow the association between age and cash holdings to be non-linear.

Married people are more likely to invest in stocks, since marriage itself is perceived as a safe asset, especially for females (Bertocchi et al., 2011; Christiansen et al., 2015). Individuals with higher educational attainment are more likely to hold riskier assets because they are likely to have higher levels of financial literacy (Calvet et al., 2007). Similarly, individuals who are better at numeracy, memory, and fluency at reading and understanding are more likely to invest in risky assets and to hold a larger proportion of risky assets in their portfolios (Atella et al., 2012; Christelis et al., 2010). Employed individuals tend to hold a higher share of risky assets compared to the unemployed (Calvet & Sodini, 2014), the retired (Viceira, 2001), and individuals working in the agriculture sector who often confront income shocks (Cocco et al., 2005). Larger households tend to have riskier portfolios, since being part of a large household implies plenty of sources of private loans, in-kind assistance, intra-household gifts and transfers, inheritances, bequests, and so on. Therefore, large households can provide a shield against income or expenditure shocks that, in other circumstances, could lead to a restructuring of portfolio choice towards a safer position (Zhang et al., 2014). However, it is also possible that larger households deal with larger budgets and expenses, which may push upward the likelihood and share of holding cash. Cardak and Wilkins (2009) find that, compared with non-homeowners, homeowners hold a larger portion of their financial wealth in risky assets. In line with this literature, we expect, married people, more educated people, people with better cognition, employed people and homeowners to be less likely to hold only cash as a safe asset and to show lower cash ratios. As for people living in larger households, they could hold either a safer or a riskier portfolio.

Peress (2004) and Guiso and Paiella (2008) show that households with large financial wealth tilt their portfolio allocations toward risky financial assets as risk aversion typically declines with wealth. In line with their findings and consistent with Karlan et al. (2014) and Prina (2015), we expect people with higher household wealth and income to be less likely to hold only cash as a safe asset and to hold a lower share of safe assets in the form of cash.

Protective health insurance schemes provide a buffer against the effect of labor income shocks and expenditure shocks following health shocks, leading to a lower need of holding safe assets (Atella et al., 2012; Goldman & Maestas, 2013). We therefore expect respondents whose insurance provides immediate reimbursement of the health care expenses to rely less on cash, as their insurance can be seen as more protective.

Karlan et al. (2014) and Prina (2015) argue that individuals living in more developed financial markets are more likely to save in banks. Hence, we expect to observe a negative association between bank accessibility and cash holdings.

Finally, the possession of mobile phones and computers as proxies for digital banking controls for the possibility that people hold their safe assets in form of digital money rather than cash.

The expected signs of all right-hand side variables in Equation (1) are illustrated in Table 1, which also provides definitions of all variables used in the paper. For our main hypothesis to hold, we expect β_1 to be positive and statistically significant.

The error term in Equation (1) includes four components. φ_p denotes a province-specific effect and is accounted for by including a full set of provincial dummies in the model. φ_t denotes a time-specific effect and is accounted for by including time dummies. φ_i represents the time-invariant individual-specific component (i.e., unobserved heterogeneity). It encompasses factors such as risk preferences, planning horizons, and/or expectations about the future that are likely simultaneously to affect both health and cash holdings. We take φ_i into account by using panel data estimators. $\varepsilon_{i,p,t}$ represents an idiosyncratic error term, which is assumed to be independent and identically distributed.

5.2. *Estimation Methodology*

We initially use a correlated random-effects (CRE) Probit model to analyze the determinants of the probability to hold only cash as a safe asset. The CRE approach was initially proposed by Mundlak (1978) for balanced panels and adapted by Wooldridge (2019) to unbalanced panels. This model is preferrable to a simple random-effects Probit model as it allows unobserved heterogeneity to be correlated with observed covariates by parameterizing the φ_i component of the error term in Equation (1) as a function of the time-averages of the time-varying variables. This boils down to estimating an augmented version of Equation (1), which includes the above-mentioned time averages using a random-effects estimator.

Focusing on the model analyzing the determinants of the proportion of safe assets held in the form of cash, because *cashratio* ranges from 0 to 1, we apply the Generalised Estimating Equation (GEE) approach proposed by Papke and Wooldridge (2008). Specifically, we estimate a fractional response model, using a Probit functional form for the mean response to impose a bounded effect of health shocks on *cashratio*. Unlike the simple fractional response model, the GEE approach accommodates panel data and allows for unobserved heterogeneity among individuals, by controlling for the within-means of time-varying variables.

6. Main Results

6.1. *Correlated Random-Effects Probit and GEE Models*

We begin our analysis by presenting correlated random-effects (CRE) Probit and GEE estimates of Equation (1) in Table 3. Column 1 reports marginal effects (MEs hereafter) of all right-hand side variables (with the exception of time and provincial dummies) in the CRE Probit models estimating the determinants of the probability to hold only cash as a safe asset. Column 3 contain the MEs in the GEE models estimating the determinants of the cash ratio. We observe that the onset of

TABLE 1
Variable definitions

Variable	Definition	
Dependent variables		
Prcash	Dummy variable (DV) equal to 1 if the respondent holds only cash as a safe asset, 0 otherwise.	
Cashratio	Ratio of cash to the sum of cash and deposits.	
Health shock indicator		Expected sign
Acute	DV equal to 1 if the respondent was diagnosed with a new acute condition in the past two years, 0 otherwise. Acute conditions include cancer, chronic lung disease, diabetes, heart disease, and stroke.	+
Control variables		Expected signs
Age	Age of the respondent.	–
Male	DV equal to 1 for males and 0 for females.	–
HH_adult2	DV equal to 1 if the respondent lives in a household with more than two adult members, 0 otherwise.	?
Married	DV equal to 1 if the respondent reports being married or cohabiting, 0 otherwise.	–
Education		
Edulow (Ref.)	DV equal to 1 if the respondent is illiterate, 0 otherwise.	N/A
Edumed	DV equal to 1 if the respondent has achieved a qualification at middle school level, 0 otherwise.	–
Eduhigh	DV equal to 1 if the respondent has at least a high school diploma, 0 otherwise.	–
Employement		
Work_emp (Ref.)	DV equal to 1 if the respondent is employed, 0 otherwise.	N/A
Work_other	DV equal to 1 if the respondent is self-employed, does agricultural work or unpaid housework, 0 otherwise.	+
Not working	DV equal to 1 if the respondent does not work (including those seeking for job and retired), 0 otherwise.	+
Cognition	Score ranging from 0 to 21 that measures the respondent's cognitive ability from the following four aspects: understanding of date/weekday/season, instant and delayed memory, numerical ability, and ability to draw a selected picture.	–
House	DV equal to 1 if the respondent's household owns at least one residential house, 0 otherwise.	–

(Continues)

TABLE 1
Continued

Variable	Definition	
Lghhitot	Logarithm of the respondent's household's real income deflated using the CPI, taking 2011 as the base year.	–
Lghfinasset	Logarithm of the respondent's household's real financial assets deflated using the CPI, taking 2011 as the base year. The financial assets mainly include cash, bank deposits, bonds, stocks, and mutual funds.	–
Mobilephone	DV equal to 1 if the respondent possesses a mobile phone, 0 otherwise.	–
Computer	DV equal to 1 if the respondent possesses a computer, 0 otherwise.	–
Bankdum	DV equal to 1 if the respondent lives in a community with at least one bank branch, 0 otherwise.	–
Reimbursedum	DV equal to 1 if the respondent participates in a health insurance scheme that provides immediate reimbursement of medical expenses, 0 if the respondent participates in a health insurance scheme that provides ex-post reimbursement.	–
Other variables used for PSM and entropy balancing		
ADL	DV equal to 1 if the respondent has any difficulty with dressing, bathing or showering, eating, getting into or out of bed, using the toilet, or controlling urination and defecation, because of health or memory problems, 0 otherwise.	N/A
IADL	Dummy equal to 1 if the respondent has any difficulty with doing household chores, preparing hot meals, shopping for groceries, managing money or taking medications, 0 otherwise.	N/A
Smoke_ever	DV equal to 1 if the respondent ever smoked, 0 otherwise.	N/A
Drink_ever	DV equal to 1 if the respondent ever drank alcohol, 0 otherwise.	N/A
SRH	Categorical variable that measures the respondent's self-reported health status: 1. Very good; 2. Good; 3. Fair; 4. Poor; 5. Very poor.	N/A
Had_minor_diseases	DV equal to 1 if the respondent was ever diagnosed with hypertension, dyslipidemia, liver disease, kidney disease, stomach or other digestive disease, emotional, nervous, or psychiatric problems, memory-related disease, asthma, arthritis or rheumatism, 0 otherwise.	N/A
Had_acute_diseases	DV equal to 1 if the respondent was ever diagnosed with an acute disease, 0 otherwise. Acute diseases include cancer, chronic lung disease, diabetes, heart disease, and stroke.	N/A

Notes: "Ref." represents the reference group in our specifications. With reference to *Bankdum*, banks typically include large commercial banks such as the Bank of China (BOC), the China Construction Bank (CCB), the Industrial and Commercial Bank of China (ICBC), the Agricultural Bank of China (ABC), and the Bank of Communications (BoCom); medium-sized and small joint-equity commercial banks; rural commercial banks; and rural credit unions (Cong et al., 2019).

TABLE 2
Descriptive statistics of outcome variables for respondents who experienced and did not experience a health shock

Acute			=0					=1			Sign. Diff
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Mean	SE	Min	Max	Obs	Mean	SE	Min	Max	Obs	
Prcash	0.673	0.469	0	1	17055	0.724	0.447	0	1	1143	***
Cashratio	0.728	0.414	0	1	17055	0.768	0.395	0	1	1143	***

Notes: Columns 1–5 (6–10) report the statistics of the outcome variables for respondents who did not experience (experienced) a health shock. Column 11 reports Wald *t*-tests for the differences in the means of the outcome variables between respondents who experienced a health shock and respondents who did not. SE denotes standard errors. Definitions of all variables are in Table 1. $^{***}p < 0.01$.

an acute condition is associated with a 3.0 percentage point (pp) higher likelihood of holding only cash as a safe asset (column 1). Similarly, the cash ratio is found to be 2.3 pp higher for people who experienced an acute health shock (column 3). These effects are economically significant: they respectively correspond to increases of 4.46 percent (3.0/0.673) and 3.16 percent (2.3/0.728) relative to the sample means of *prcash* and *cashratio* reported in Table 2 for respondents not experiencing the health shock.

From a quantitative viewpoint, these effects compare favorably with the 1.69 pp lower probability of participating in the stock market and the 0.35 pp lower fraction of financial wealth invested in risky assets identified by Døskeland and Kvaerner (2022) as a result of a cancer diagnosis in Norway. They are also in line with the 2.1, 0.2, and 1.7 pp lower probabilities of owning retirement accounts, bonds, and risky assets among US middle-aged and older people experiencing ill health highlighted by Rosen and Wu (2004).

These results are consistent with our main hypothesis and with the literature which reveals that a poor health status and health shocks are associated with safer portfolios (Atella et al., 2012; Berkowitz & Qiu, 2006; Døskeland & Kvaerner, 2022; Rosen & Wu, 2004).

Focusing on the control variables, in line with the literature described in Section 5.1, we find that being male, having higher levels of financial wealth, and living in a community with banks are all associated with a lower likelihood of holding only cash as a safe asset, and with a lower cash ratio. By contrast, being married is positively associated with cash holdings. Although this finding is in contrast with Bertocchi et al. (2011) and Christiansen et al. (2015), who argue that married people generally hold riskier portfolios, it can be rationalized bearing in mind that in the context of China, married people have been found to have less trust in banks (Fungáčová & Weill, 2018).

6.2. *Verifying Whether the Results are Robust to Using a Linear Fixed-Effects Estimator*

In columns 2 and 4 of Table 3, we use a linear fixed-effects (FE) estimator instead of the CRE Probit estimator (for *prcash*) and the GEE estimator (for

TABLE 3
BASELINE SPECIFICATIONS

Dep Var	Prcash		Cash ratio	
	(1) CRE Probit	(2) Linear FE	(3) GEE	(4) Linear FE
Acute	0.030***	0.051**	0.023**	0.033*
	(0.011)	(0.022)	(0.010)	(0.018)
Age	−0.002	−0.015	−0.005	−0.013
	(0.009)	(0.016)	(0.007)	(0.013)
Age^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Male	−0.055***	-	−0.049***	-
	(0.007)	-	(0.006)	-
Married	0.097***	0.095**	0.086***	0.055
	(0.023)	(0.046)	(0.025)	(0.036)
Edu_med	0.018	0.015	0.012	0.004
	(0.018)	(0.032)	(0.017)	(0.026)
Edu_high	−0.034	−0.048	−0.018	−0.049
	(0.037)	(0.068)	(0.029)	(0.061)
Work_other	−0.001	0.005	−0.005	0.002
	(0.013)	(0.023)	(0.011)	(0.019)
Not working	−0.005	0.004	−0.005	0.000
	(0.010)	(0.018)	(0.008)	(0.015)
Cognition	−0.001	−0.001	−0.001	−0.000
	(0.001)	(0.002)	(0.001)	(0.002)
HH_adult2	0.001	0.002	−0.000	−0.001
	(0.010)	(0.018)	(0.008)	(0.015)
Lghhitot	0.002	0.002	0.002	0.001
	(0.001)	(0.002)	(0.001)	(0.002)
Lghfinasset	−0.118***	−0.138***	−0.109***	−0.126***
	(0.001)	(0.004)	(0.002)	(0.003)
House	−0.011	−0.011	−0.008	−0.015
	(0.011)	(0.019)	(0.011)	(0.016)
Computer	0.009	0.015	0.016*	0.024
	(0.010)	(0.019)	(0.009)	(0.016)
Mobilephone	−0.014	−0.003	−0.014	−0.006
	(0.013)	(0.023)	(0.012)	(0.018)
Bankdum	−0.056***	-	−0.047***	-
	(0.010)	-	(0.008)	-
Reimbursedum	0.010	−0.071	0.003	−0.006
	(0.023)	(0.097)	(0.019)	(0.017)
Province & Year FE	Yes	Yes	Yes	Yes
N	18198	18198	18198	18198

Notes: This table reports marginal effects obtained using a CRE Probit model (column 1), a GEE model (column 3), and a linear fixed-effects model (columns 2 and 4). In columns 1 and 2 (3 and 4), the dependent variable is the probability of holding only cash as a safe asset (the cash ratio). Robust standard errors (standard errors clustered at the household-level) are reported in parentheses in columns 1 and 3 (2 and 4). Definitions of all variables are in Table 1. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

c*ashratio*). The advantage of FE models is that they allow for the correlation between unobserved heterogeneity and time-varying covariates. However, a key drawback to such models is that time-invariant factors, such as gender and bank accessibility in our case, are not estimable. In line with previous results, we observe a 5.1 pp higher likelihood of holding only cash as a safe asset (column 2), and a 3.3 pp higher cash ratio for those respondents who experienced the onset of an acute condition (column 4). The coefficients associated with the control variables are similar to those reported in columns 1 and 3.

6.3. *Accounting for Attrition*

Sample attrition is likely to be present in our dataset and may affect our results. Specifically, given the longitudinal nature of our dataset, individuals drop out from the panel at each wave and some of these drop-outs may not be random, but instead related to risk preference (which affects portfolio choice) and health status (e.g., death).

To check whether our empirical results are affected by attrition bias, we adopt the following two variable addition tests initially suggested by Verbeek and Nijman (1992, p. 688). Statistical significance of the added variable in the baseline specification provides a test for attrition bias. First, we verify whether our CRE Probit (GEE) estimates for the probability of holding only cash as a safe asset (the cash ratio) are robust to including among the covariates the number of waves each respondent took part in. The results, which are not reported for brevity, but are available upon request, show that the marginal effects associated with this new variable are insignificant. Moreover, the sign and significance of the marginal effects associated with the health shock (as well as those of the control variables) do not change after including this new variable.

Second, we check whether our results are robust to including a dummy variable equal to 1 if the respondent is present in all three waves of the survey, and 0 otherwise. The results, which are not reported for brevity, show that, once again, the marginal effects associated with this new variable are not statistically significant. Moreover, the marginal effects associated with the health shock variables (and other controls) do not change when this indicator is included. Both sets of results suggest that attrition bias is not present, and that attrition does not distort the marginal effects of the key and control variables included in our models.

Finally, to mitigate the effects of attrition, Banks et al. (2010) estimate their models on a balanced panel. Using a balanced panel ensures that different trajectories observed between groups are not due to differential attrition. Our results were robust to restricting our sample to a balanced panel. This suggests that the results in Table 3 are not driven by attrition.

7. Establishing a Causal Relationship Between Health Shocks and Cash Holdings

The key challenge of evaluating the causal link between health shocks and cash holdings is to ensure any changes in cash holdings are due to the health shock and would not have occurred without that shock. However, in real life, it is impossible to observe the outcome in the absence of the event. Hence, the results obtained in our baseline specifications could be driven by systematic differences between individuals experiencing a health shock and individuals who do not experience the shock. We tackle this problem in three ways: using an event study framework as in Dobkin et al. (2018), using a propensity score matching approach, and using an entropy balancing approach. These three methods and the results we obtain using them are discussed in turn hereafter.

7.1. *Event Study*

In order to establish a causal relationship between health shocks and cash holdings, we begin by making use of a non-parametric event study approach, which boils down to estimating a general model of the following type (Dobkin et al., 2018):

$$\begin{aligned} Y_{i,p,t} = {} & \beta_{11} Period(-2) + \beta_{12} Period(0) + \beta_{13} Period(1) + \beta_{14} Period(2) \\ & + \gamma' X_{i,p,t} + \delta' Z_{i,p} + \text{error term } (i = 1, \dots, \text{N};\ t = 1, 2, 3;\ \text{p} = 1, \dots, 24) \end{aligned} \tag{2}$$

where i denotes respondents; t, time; and p, provinces. $\boldsymbol{Y}$ represents in turn one of the two outcome variables we consider, namely *prcash* and *cashratio*. $\boldsymbol{X}$ and $\boldsymbol{Z}$ respectively denote the same time-varying and time-invariant control variables included in Equation (1). *Period*(*−2*) indicates the second wave preceding the survey wave in which the onset of the health shock (event) is reported to have occurred in the last 2 years. *Period*(*0*) indicates the wave in which the event is reported to have occurred in the last 2 years, and *Period*(*1*) and *Period*(*2*) indicate in turn the first and second wave following the wave in which the event is reported to have occurred in the last 2 years. As in Dobkin et al. (2018), *Period*(*−1*) is the omitted category. The structure of the error term is the same as in Equation (1). In this framework, we aim at analyzing the coefficients associated with the indicator variables for time relative to the event. The primary advantage of this non-parametric event study is that it allows us to assess the pattern of *prcash* and *cashratio* relative to *Period*(*−1*).

To interpret the coefficients β_{12}–β_{14} as the causal effect of the health shock on respondents' portfolio choice, the identifying assumption is that, conditional on the onset of an acute health condition and the included control variables, the timing of the health shock is uncorrelated with the portfolio choice of the respondents, which is a plausible assumption (Dobkin et al., 2018). Moreover, our specification enables us to verify this by estimating the pre-event outcomes.

Estimates of Equation (2) are presented in columns 1 and 2 of Table 4. In column 1, the dependent variable is the probability of holding only cash as a safe asset and a CRE Probit model is used in estimation. In column 2, the dependent variable is the cash ratio, and a GEE estimator is used in estimation. As only respondents who actually suffered from a health shock over the sample period are considered in the estimation of Equation (2), the sample used in Table 4 is small (4045 observations). We can see that only the coefficient associated with the *Period*(*0*) dummy is statistically significant.

Columns 3 and 4 of Table 4 report estimates of a parametric version of the event study model shown in Equation (2). Following Dobkin et al. (2018), this consists in replacing the *Period*(−*t*) dummies with a linear trend in the number of waves

TABLE 4
Event study

	Non-parametric event study		Parametric event study	
Dep Var	Prcash	Cashratio	Prcash	Cashratio
	(1) CRE Probit	(2) GEE	(3) CRE Probit	(4) GEE
Period (−2)	0.017	0.012		
	(0.017)	(0.013)		
Period (0)	0.050***	0.030**	0.050**	0.041*
	(0.016)	(0.013)	(0.025)	(0.021)
Period (1)	0.009	−0.002	0.011	0.011
	(0.020)	(0.017)	(0.030)	(0.025)
Period (2)	0.023	0.014	0.024	0.026
	(0.028)	(0.024)	(0.036)	(0.030)
Linear pre-trend			0.003	0.009
			(0.012)	(0.010)
Age	0.011	0.011	0.010	0.010
	(0.017)	(0.014)	(0.017)	(0.014)
Age^2	−0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Male	−0.036***	−0.034***	−0.036***	−0.034***
	(0.013)	(0.010)	(0.013)	(0.010)
Married	0.144***	0.123***	0.144***	0.123***
	(0.042)	(0.044)	(0.043)	(0.044)
Edu_low	−0.034	−0.021	−0.033	−0.020
	(0.075)	(0.053)	(0.075)	(0.053)
Edu_med	0.021	−0.003	0.022	−0.002
	(0.069)	(0.046)	(0.069)	(0.046)
Work_other	−0.039	−0.046**	−0.039	−0.044*
	(0.028)	(0.023)	(0.028)	(0.023)
Not working	−0.027	−0.037*	−0.027	−0.036*
	(0.024)	(0.019)	(0.024)	(0.019)
Cognition	0.001	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
HH_adult2	0.014	0.006	0.013	0.005
	(0.018)	(0.015)	(0.018)	(0.015)
Lghhitot	0.002	0.001	0.002	0.001
	(0.003)	(0.002)	(0.003)	(0.002)
Lghfinasset	−0.107***	−0.097***	−0.107***	−0.097***
	(0.003)	(0.003)	(0.003)	(0.003)
House	−0.030	−0.030	−0.030	−0.030
	(0.023)	(0.020)	(0.023)	(0.020)
Computer	0.002	0.009	0.002	0.008
	(0.022)	(0.018)	(0.022)	(0.018)
Mobilephone	0.012	0.004	0.012	0.004
	(0.024)	(0.019)	(0.024)	(0.019)
Bankdum	−0.050**	−0.039**	−0.050**	−0.039**
	(0.021)	(0.016)	(0.021)	(0.016)
Reimbursedum	−0.036	−0.028	−0.036	−0.028
	(0.025)	(0.021)	(0.025)	(0.021)
Province & Year FE	Yes	Yes	Yes	Yes
N	4045	4045	4045	4045

Notes: In columns 1 and 3 (2 and 4), the dependent variable is the probability of holding only cash as a safe asset (the cash ratio). Columns 1 and 3 (2 and 4) report marginal effects obtained using a CRE Probit model (GEE model). *Period(r)* refers to the survey wave relative to the survey wave in which the onset of an acute health shock is reported to have occurred in the last 2 years ($r=0$). *Linear pre-trend* is a linear trend in the number of waves preceding the wave in which the onset of the health shock (event) is reported to have occurred in the last 2 years. Robust standard errors are reported in parentheses. Definitions of all variables are in Table 1. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

preceding the wave in which the onset of the health shock (event) is reported to have occurred in the last 2 years (also see Miller et al., 2021, for a similar approach). In this case, the coefficients associated with *Period*(*0*), *Period*(*1*), and *Period*(*2*) indicate changes in cash holdings following the onset of an acute condition relative to any pre-existing trend.[3] In both columns 3 and 4 of Table 4, we can see that the coefficients associated with the linear pre-trend are not statistically significant. Moreover, as in columns 1 and 2, only the coefficient *Period*(*0*) is statistically significant.

In a nutshell, these results suggest that the effects of the health shock on cash holdings only appear in the wave in which the onset of the health shock is reported to have occurred in the last 2 years. This can be explained by recognizing that the CHARLS is a bi-annual survey. As a result, if the effects of the health shock are short-lived, people are likely to revert to the pre-shock portfolio allocation more than 2 years after the onset of the shock. In the case in which the effects of the health shock are long-lasting, people adjust to the new normal and learn to manage their medical expenditures (and associated reimbursement procedures) more efficiently. In columns 1 and 2 of Table 4, we also observe that the coefficients associated with the *Period*(−*2*) dummy are not statistically significant. Together with the fact that the coefficient associated with the linear pre-trend in columns 3 and 4 is also insignificant, this confirms that the timing of the health shock is uncorrelated with cash holdings.

7.2. *Propensity Score Matching: Average Treatment Effects on the Treated and Estimates Conducted on a Matched Sample*

Average Treatment Effects on the Treated

To establish a causal relationship between health shocks and cash holdings, we hereafter make use of a Propensity Score Matching (PSM) technique (Rosenbaum & Rubin, 1983). This consists in matching individuals who have experienced a given health shock (treatment group) with other individuals who have not experienced the shock but are as similar as possible in many other aspects to the treatment group (control group) and comparing the outcomes in terms of cash holdings between the two groups.

Specifically, we start by estimating a Probit model to gauge the probability of experiencing a health shock (i.e., the probability of being "treated") as a function of all the variables used in our cash holdings models, as well as the following new variables measuring the individual's health status and risk factors: a dummy variable (DV) equal to 1 if the individual suffers from any limitations in activities of daily living (i.e., dressing, walking across a room, bathing or showering, eating, getting in or out of bed, and using the toilet), and 0 otherwise; a DV equal to 1 if the individual suffers from any limitations in instrumental activities of daily living (i.e., housework, shopping, cooking, managing money, and taking medicine), and

[3] Compared to the non-parametric specification, the identifying assumption in the parametric specification is weaker and requires that, conditional on the onset of an acute health shock and the included control variables, the timing of the health shock is uncorrelated with deviations in cash holdings from a linear trend in event time.

0 otherwise; the respondent's self-assessed health; a DV equal to 1 if the respondent ever suffered from an acute health condition (i.e., cancer, chronic lung disease; diabetes; heart disease; and stroke) in the past, and 0 otherwise; a DV equal to 1 if the respondent ever suffered from a mild health condition (e.g., hypertension; dyslipidemia; high blood sugar; liver disease; kidney disease; emotional, nervous, or psychiatric problems; memory-related disease; arthritis or rheumatism; asthma) in the past, and 0 otherwise; a DV equal to 1 if the respondent ever smoked, and 0 otherwise; and a DV equal to 1 if the respondent ever drank alcohol, and 0 otherwise.[4] In so doing, we follow Stuart (2010), according to which variables that may be associated with treatment assignment and/or outcome should be included in the model aimed at estimating the propensity scores. In order to avoid post-treatment bias, all explanatory variables in the Probit model (with the exception of gender) are lagged one wave. The marginal effects obtained from the estimation of this Probit model are reported in Appendix Table S2. Fitted values from this regression give the propensity score.

Following Rosenbaum and Rubin (1983), we then pair each individual from the treatment group with an individual from the control group using the nearest-neighbor criterion (i.e., the closest propensity score) with sample replacement and caliper set at 0.01.[5] In order to improve matching quality, we impose the common support, which helps avoid matching bias by dropping those treated observations whose propensity scores are higher than the maximum or lower than the minimum of the propensity score of untreated respondents.

We define the average treatment effect of experiencing a health shock on the cash holdings of respondents who actually experienced the health shock (i.e., the "treated" respondents) as follows:

$$\text{ATT} = E\left(Y_1 - Y_0|D=1\right) = E\left(Y_1|D=1\right) - E\left(Y_0|D=1\right) \tag{3}$$

where $\boldsymbol{Y_1}$ represents the cash holdings of a respondent who experienced the health shock, and $\boldsymbol{Y_0}$, the cash holdings of a respondent who did not. $\boldsymbol{D}=1$ denotes the treatment status (i.e., experiencing the health shock). At the individual level, the first term on the right-hand side of Equation (3) represents the cash holdings of subject $\boldsymbol{X}$ who experienced a health shock. The second term represents the counterfactual: it measures what the cash holdings status of a person who experienced a health shock would have been had he/she not experienced the shock. As this counterfactual is unobservable for subject $\boldsymbol{X}$, we seek an alternative subject, $\boldsymbol{Z}$ (taken from the control group), with the same characteristics as subject $\boldsymbol{X}$, and observe their cash holdings in the absence of the health shock. In other words, we use this as a surrogate outcome for subject $\boldsymbol{X}$'s counterfactual outcome. Extending this to a group level enables us to calculate the average treatment effect on the treated (ATT).

The ATTs are reported in Table 5. In line with our baseline results and with our main hypothesis, they suggest that experiencing the onset of an acute condition

[4] Definitions of these variables are presented in Table 1.

[5] Dehejia and Wahba (2002) show that matching with replacement reduces bias compared to matching without.

TABLE 5
AVERAGE TREATMENT EFFECTS ON THE TREATED (PSM)

Outcome	ATT	Std. err.	T-stat	$P > T$	Total obs.
Prcash	0.045***	0.020	2.21	0.014	16689
Cashratio	0.032***	0.018	1.78	0.075	16689

Notes: This table reports the average treatment effects on the treated (ATTs) and relevant statistics obtained using PSM. The treatment considered is *Acute*. We use a one-to-one nearest neighbor matching algorithm with replacement and caliper of 0.01 for individuals in the area of common support. Definitions of all variables are in Table 1. ***$p < 0.01$.

results in a 4.5 pp higher chance of holding only cash as a safe asset, and in a 3.2 pp higher cash ratio.

Assessing the Quality of the Matching

To gauge the quality of the matching procedure, we run variable-specific balancing tests (Leuven & Sianesi, 2003). Appendix Table S3 presents the results of these tests. Specifically, the table reports the means of all variables used to obtain the propensity score for the treated and control groups in the matched sample, indicating whether the differences in these means across the two groups are statistically significant. We can see that despite the relatively large size of our sample, for all our conditioning variables, the null hypothesis of no difference in means between treated and matched controls after matching cannot be rejected. This is reassuring as the t-tests are heavily dependent on sample size (Imbens & Wooldridge, 2009). The table also shows that the mean variable-specific standardized percentage bias after matching is generally lower than the 5 percent threshold recommended by Rosenbaum and Rubin (1985). Finally, Rubin (2001) recommends Rubin's B-statistic to be less than 25, and Rubin's R-statistic to be between 0.5 and 2 for the samples to be sufficiently balanced, which is what we find. Taken together, these statistics suggest that the quality of our matching is good. In other words, the matching process excludes meaningful differences along observed matching dimensions between respondents in the treatment and control groups.

Estimation of the Baseline Model on a Matched Sample

Next, to minimize the effects of selection based on observed characteristics in our baseline model, we use the PSM method described above to create a matched sample of treated and control individuals which are as similar as possible. We then re-estimate Equation (1) on this highly comparable propensity score-matched sample.

Appendix Table S4 presents the estimates of our baseline models conducted on the matched sample. In line with the results in Table 3, respondents who experienced the onset of an acute condition are 4.3 pp more likely to hold only cash as a safe asset (column 1). Furthermore, the cash ratios for rural residents who experienced the onset of an acute condition are on average 3.0 pp higher than those of respondents who did not experience any health shock (column 2).

7.3. *Entropy Balancing: Baseline and GMM Regressions*

Entropy Balancing: Baseline Models

As discussed above, differences in the covariate distributions between people subject and not subject to acute health shocks may confound any identification of the link between health shocks and cash holdings. Entropy balancing eliminates these differences by enforcing covariate balancing in a constrained, nonlinear estimation approach. Specifically, entropy balancing is implemented in two consecutive steps. First, observations are re-weighted with respect to the treatment (i.e., experiencing a health shock) so that all the relevant matching covariates are balanced (i.e., they have very similar mean and variance). Through this re-weighting procedure, a region of common support is generated, where respondents experiencing and not experiencing a health shock are comparable on the matching covariates. In other words, a synthetic control group as similar as possible to the treatment group is created. Second, the weights obtained in the first step are used in a regression analysis with the treatment indicator (i.e., experiencing a health shock) as an explanatory variable.

The entropy balancing approach overcomes several drawbacks of the PSM method. First, contrary to the PSM, it ensures that higher-order moments of the covariate distributions are nearly identical across treated and synthetic control samples. Ensuring covariate balance increases the plausibility that any differences in cash holdings we document are driven by the treatment rather than correlated differences in determinants. Second, contrary to PSM, entropy balancing is nonparametric in the sense that no empirical model for the selection into treatment needs to be specified. Consequently, the misspecification of the functional form of the empirical model and associated bias is ruled out. In other words, statistical inferences obtained through entropy balancing are less sensitive to design choices. Third, because it is in essence a weighted regression, using entropy balancing means that, contrary to the PSM case whereby each observation in the treatment group is matched with only one (or few) observations in the control group, all observations in the sample can be used. By retaining information in the control sample, this approach maximizes estimation power (Watson & Elliot, 2016). Fourth, using Monte Carlo simulations as well as empirical applications, Hainmueller (2012) demonstrates that entropy balancing outperforms PSM in terms of estimation bias and mean square error. In other words, our treatment comes closer to randomization since we obtain a much higher degree of covariate balance. Finally, unlike conventional matching, entropy balancing allows us to exploit fully the panel nature of our dataset in the second stage. As a result of these advantages, entropy balancing is frequently employed in recent studies in health economics, international economics, and finance (e.g., Apeti & Edoh, 2023; Chen et al., 2022; Egger et al., 2020; Shoji et al., 2022).

We balance relevant observable characteristics on two moments (mean and variance) with a tolerance of 0.015. These characteristics are all the covariates included in our models as well as the following variables (lagged): self-reported health indicator, dummies for whether the respondent suffers from limitations in activities of daily living or activities instrumental to daily living, dummies for whether the respondent ever smoked or drank alcohol, and for whether he/she ever suffered from mild or severe health conditions. Appendix Table S5 presents the mean and variance of relevant matching characteristics in the treatment and synthetic control group before and after weighting by the entropy balancing weights. While we see large distributional differences between the two groups for almost all matching covariates before balancing, the first and second moments of the covariate distributions are virtually identical among respondents experiencing and not experiencing a health shock after entropy balancing. In addition, Appendix Figure S1 depicts the histogram of the entropy balancing weights for experiencing an acute health shock. They mostly range from zero to one, suggesting that the covariate balance is achieved without placing too large weights on a few observations.

The second stage results are presented in columns 1 and 2 of Table 6.[6] In line with previous findings, we can see that the onset of an acute health condition leads to a 5.2 pp higher probability of holding only cash as a safe asset, as well as a 3.5 pp higher cash ratio.[7]

Entropy Balancing: Using a Dynamic System GMM Estimator

To account for the persistence in cash holding over time, we estimate a dynamic version of Equation (1) using the system GMM estimator, which jointly estimates the equation in first differences and levels (Arellano & Bover, 1995; Blundell & Bond, 1998). By estimating the model in first-differences, we are able to control for time-invariant individual-specific characteristics that might affect both the probability of experiencing a health shock and cash holdings. In doing so, the risk of confounding is considerably reduced. In addition to controlling for unobserved heterogeneity, the GMM estimator overcomes the possible endogeneity of regressors, which may cause simultaneity bias. The lagged dependent variable and any other potentially endogenous variables are typically instrumented with lags of themselves (Roodman, 2009).

Yet, estimating our baseline specification in first-differences may lead to three shortcomings. First, if the original model is conceptually in levels, differencing may reduce the variation in the explanatory variables (Beck et al., 2000). Second, first-differencing may exacerbate the impact of measurement error on the dependent variables (Griliches & Hausman, 1986). Third, variables in levels may be weak instruments for first-differenced equations (Arellano & Bover, 1995). To take these shortcomings into account, Arellano and Bover (1995) and Blundell and Bond (1998) suggest estimating the equation in first differences jointly with the

[6]It should be noted that given that weighted regressions cannot be conducted for correlated random-effects Probit or GEE models, these results are all obtained using a fixed-effects linear model. However, in Table 3, we have shown that in the baseline model, using a fixed-effects linear model delivers results very similar to those obtained using the correlated random-effects Probit model or the GEE model.

[7]The smaller number of observations in columns 1 and 2 of Table 6 compared to Table 3 can be explained by the missingness of some of the matching variables used in the entropy balancing procedure, as well as by the fact that singletons are dropped automatically from the linear fixed-effects regressions generated using the *regdhfe* command in Stata. In our models, we account for individual, province, and time fixed effects. Singletons are groups with only one observation. Correira (2015) document that singletons can overstate statistical significance and lead to incorrect inference.

TABLE 6
BASELINE SPECIFICATIONS ESTIMATED USING ENTROPY BALANCING: FIXED-EFFECTS AND GMM MODELS

Dep Var	Prcash	Cashratio	Prcash	Cashratio
	(1) FE	(2) FE	(3) GMM	(4) GMM
Acute	0.052***	0.035**	0.067**	0.045*
	(0.019)	(0.014)	(0.028)	(0.024)
Age	−0.026	−0.025	−0.002	−0.003
	(0.021)	(0.016)	(0.013)	(0.011)
Age^2	0.0001	0.00017	−0.000	−0.000
	(0.0002)	(0.00012)	(0.000)	(0.000)
Married	0.077	0.053	−0.059***	−0.051***
	(0.083)	(0.057)	(0.016)	(0.014)
Edu_med	0.053	0.010	0.035	0.052
	(0.045)	(0.032)	(0.061)	(0.050)
Edu_high	−0.065	−0.094	−0.018	−0.021
	(0.068)	(0.066)	(0.021)	(0.018)
Work_other	0.025	0.024	−0.063	−0.044
	(0.029)	(0.023)	(0.043)	(0.034)
Not working	0.023	0.012	−0.014	−0.055
	(0.021)	(0.017)	(0.055)	(0.044)
Cognition	−0.003	−0.000	−0.005	−0.045
	(0.003)	(0.002)	(0.046)	(0.037)
HH_adult2	0.029	0.020	−0.008**	−0.004
	(0.020)	(0.017)	(0.004)	(0.003)
Lghhitot	0.007**	0.004	−0.009	0.008
	(0.003)	(0.003)	(0.031)	(0.026)
Lghfinasset	−0.131***	−0.120***	0.012	−0.001
	(0.005)	(0.004)	(0.028)	(0.022)
House	−0.033*	−0.040***	−0.103**	−0.126***
	(0.019)	(0.016)	(0.042)	(0.037)
Computer	−0.016	−0.001	0.003	0.003
	(0.025)	(0.018)	(0.024)	(0.020)
Mobilephone	0.021	0.016	0.015	0.024
	(0.033)	(0.027)	(0.036)	(0.029)
Reimbursedum	−0.206	−0.132	0.039	0.044**
	(0.145)	(0.085)	(0.025)	(0.021)
Bankdum			−0.065**	−0.044*
			(0.031)	(0.026)
Lagged prcash			0.096**	
			(0.039)	
Lagged cashratio				0.086*
				(0.045)
Province & Year FE	Yes	Yes	Yes	Yes
Hansen-J test (*p-value*)			0.242	0.090
N	13244	13244	11497	11497

Notes: In column 1 and 3 (2 and 4), the dependent variable is the probability of holding only cash as a safe asset (the cash ratio). Columns 1 and 2 report coefficients obtained using a weighted linear fixed-effects model, with weights derived from the entropy balancing approach. Columns 3 and 4 report system GMM estimates weighted using weights obtained from the entropy balancing approach. In columns 3 and 4, we treat lagged *prcash*/*cashratio*, household income, household financial wealth, and the health shock as potentially endogenous variables. We instrument the former three variables with their own lags, and the health shock with the following lagged variables: self-perceived health, smoking and drinking behaviors, suffering from limitations in activities of daily living or activities instrumental to daily living, and dummies indicating whether the respondent ever suffered from mild or severe health shocks in the past. Levels of all these variables dated $t-2/t-3$ are used as instruments in the first-differenced equations, and the first differences of these same variables lagged once/twice are used as additional instruments in the levels equations. The Hansen-J test of over-identifying restrictions is distributed as Chi-square under the null of instrument validity. Standard errors clustered at household level (robust standard errors) are reported in parentheses in columns 1 and 2 (3 and 4). Definitions of all variables are in Table 1. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

same equation in levels. First-differenced variables can be used as instruments for the equations in levels in a system of equations that includes the equations in both levels and differences.

In summary, the system GMM estimator enables us to obtain efficient estimates while controlling for time-invariant unobserved heterogeneity and simultaneity. It also allows us to consider the dynamic nature of the relationship between health shocks and cash holdings. An additional important advantage of the system GMM estimator is that it enables us to estimate coefficients on time-invariant variables such as gender. For datasets that consist of large cross sections with a small time-series dimension (such as ours), the system GMM is found to be more efficient than the simple first-difference GMM estimator (Blundell & Bond, 1998; Roodman, 2009).

Columns 3 and 4 of Table 6 present estimates of a dynamic version of Equation (1) obtained using the system GMM estimator weighted using weights obtained from the entropy balancing approach. Using these weights enables us to also take into account the selection bias which arises if health shocks and cash holdings both depend on the same common variables. Column 3 (4) presents the estimates of the model in which the dependent variable is the probability of holding only cash as a safe asset (the cash ratio). In all specifications, in addition to the lagged dependent variable, we also treat household income, financial wealth, and the health shock as endogenous. Levels of *prcash/cashratio*, income, and wealth dated $t-2$ and/or $t-3$ are used as instruments in the first-differenced equations. The health shock is instrumented with two and/or three lags of self-perceived health, the dummies indicating smoking and drinking behaviors, the dummies indicating limitations in activities of daily living or activities instrumental to daily living, and the dummies indicating whether the respondent ever suffered from mild or severe health shocks in the past. First differences of these variables lagged once/twice are used as additional instruments in the levels equations.

Focusing on *prcash*, the coefficient associated with the onset of an acute condition is significant and positive, implying a 6.7 pp higher probability of holding only cash as a safe asset in the presence of an acute health shock (column 3). The same is found for the cash ratio: an acute health shock is associated with a 4.5 pp higher cash ratio (column 4). These findings are in line with our baseline results.

To ensure the validity of our instruments and the specification of the models, we present the *Hansen* (J) test. In all specifications, this test does not show any problems with instrument choice or model specification.[8]

8. The Role of Reimbursement Method and Bank Accessibility

We next explore the role of reimbursement method and bank accessibility as potential drivers of the association we found between health shocks and cash holdings.

[8] The tests for second- and third-order serial correlation of the differenced residuals are not reported as we do not have a sufficient number of time-series observations to compute them. Also note that the smaller number of observations in columns 3 and 4 of Table 6 compared to Table 3 is due to the presence of the lagged dependent variable.

8.1. *Reimbursement Method*

Motivation

Various studies show that protective health insurance can mitigate the effect of health shocks on portfolio choice (Goldman & Maestas, 2013; Lee, 2018). Yet, although almost all residents in rural China are covered by the NRCMS, the consequences of health shocks can be particularly severe due to the relatively low reimbursement rate of the NRCMS.

Moreover, although according to the National Health and Family Planning Commission, P.R. China (2014), by 2014, the NRCMS covered about 98.7 percent of the total rural population, in our sample, just under 70 percent of the NRCMS beneficiaries have to pay the full medical cost out of their own pockets before receiving a reimbursement ex-post, while other beneficiaries get immediate reimbursement (i.e., they only pay at the coinsurance rate for each treatment exceeding the deductible).

Both the level of reimbursement and whether the respondent gets the reimbursement immediately or later are not a function of household/individual characteristics nor of the contributions paid. As Zhou et al. (2016) put it: "Individuals cannot choose the insurance type, but rather must accept the cost-sharing method offered by the insurance and local government" (p. 42).[9] These authors further argue that: "In China, health insurance and cost-sharing methods are not self-selected but rather decided by local policies" (p. 57). China's current medical insurance systems are mainly coordinated at the county-level, driving significant geographical variation on reimbursement rates, coverage depth and cost-sharing policy (Liu et al., 2017; Zhang & Zhang, 2022; Zhou et al., 2016). In line with this argument, Meng et al. (2015) document that NCMS funds are pooled at the county level, which implies that in China there are roughly 2852 NCMS schemes. They further confirm that the benefit packages and financial protection are not equal within and across the schemes. As a result, policies such as coverage and cost-sharing usually depend on the economic performance of the counties, rather than the contribution of the beneficiaries (Yin et al., 2019).

We believe that whether or not the respondent benefits from the immediate reimbursement of medical expenses (which in light of the arguments discussed above can be considered an exogenous factor) can significantly affect the association between health shocks and cash holdings. As rural areas are typically characterized by a low level of financial development, people facing health shocks and not benefitting from immediate reimbursement of their healthcare expenditures will prefer to hold more cash in order to face the uncertain payments associated with their new illness. The uncertainties surrounding the amount and timing of the

[9] Zhou et al. (2016) explain that the co-payment program requires participants to pay only their shares of the medical expenses at the time of service while the service provider applies for payment from the insurance program after services are rendered. In our paper, we refer to this scheme as "immediate reimbursement". Zhou et al. (2016) further explain that, in the case of delayed reimbursement, the reimbursement program requires participants to pay the whole cost of medical services at the time of service. Then, participants apply for reimbursement after services are rendered through their registered health bureaus.

reimbursement (Zhong, 2011) will provide a further incentive to hold cash, which is the safest of safe assets. By contrast, respondents whose insurance provides immediate reimbursement will face lower healthcare expenditures following a health shock, coupled with lower uncertainty surrounding future medical expenditures and reimbursements. As a result, they will make smaller or no changes to their portfolio.

Testing

To test whether ex-post reimbursement of medical expenses drives the association between health shocks and cash holdings, we estimate separate models for respondents who benefit and do not benefit from immediate reimbursement of their health expenditures. We expect the association between health shocks and cash holdings to be stronger for the latter.

Main Results

Panel A of Table 7 reports estimates of the coefficients associated with *Acute* in Equation (1) for respondents with and without immediate reimbursement of the health expenditures. In column 1, the dependent variable is *prcash*, whilst in column 2, it is *cashratio*. All estimates are carried out using a fixed-effects estimator, weighted using the weights obtained from the entropy balancing approach. As expected, the coefficient associated with the health shock is only statistically significant for individuals who do not benefit from immediate reimbursement of medical expenses. Specifically, for these respondents, the occurrence of an acute health shock is associated with a 7.2 pp higher likelihood of holding only cash as a safe asset (column 1), and with a 5.4 pp higher cash ratio (column 2).

The absence of a significant association between health shocks and cash holdings for people who benefit from immediate reimbursement of medical expenses is consistent with Goldman and Maestas (2013), who find that protective health insurance mitigates the effect of health shocks on portfolio choice, and with Angrisani et al. (2018) and Lee (2018), who document that people with protective health insurance hold riskier portfolios.

Additional Results

We next exploit the geographical variation in the timing of the reimbursement of the health expenditures as follows. First, we calculate the median of the reimbursement dummy by community and wave. Second, we differentiate communities into those with average value of the reimbursement dummy higher and lower than the median in each year. Third, we estimate relevant cash holding models differentiating respondents based on whether they reside in communities with high or low average values of the reimbursement dummy in each wave. We expect the latter to show a stronger sensitivity of cash holdings to the onset of an acute health condition. The results, which are presented in Panel B of Table 7 confirm that it is only the respondents who live in communities where the immediate reimbursement of medical expenses is less widespread who show a significant association between health shocks and cash holdings.

TABLE 7
The role of reimbursement method and bank accessibility (entropy balancing approach)

Dep Var	Prcash	Cash ratio	*N*
	(1)	(2)	(3)
Panel A: Reimbursement type			
Ex-post	0.072***	0.054***	9016
	(0.025)	(0.021)	
Immediate	0.015	−0.001	4228
	(0.053)	(0.038)	
Panel B: Prevalence of immediate reimbursement (community-level)			
Lower than the median	0.107***	0.091***	6729
	(0.032)	(0.027)	
Higher than the median	0.001	−0.012	6515
	(0.040)	(0.029)	
Panel C: Bank dummy			
Communities without banks	0.049**	0.034**	11999
	(0.026)	(0.020)	
Communities with banks	0.086	0.043	1245
	(0.068)	(0.065)	
Panel D: Number of banks per 1000 people (city-level)			
Lower than median	0.072**	0.048*	5842
	(0.035)	(0.028)	
Higher than median	0.051	0.040	5979
	(0.037)	(0.030)	

Notes: This table reports the coefficients associated with the *Acute* dummy in regressions estimated on different subsamples of respondents. Estimates are obtained using a weighted fixed-effects linear model, with weights derived from the entropy balancing approach. The dependent variable is the probability of holding only cash as a safe asset (the cash ratio) in columns 1 (2). Column 3 reports the size of each sub-sample. Standard errors clustered at household level are reported in parentheses. Definitions of all variables are in Table 1. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

8.2. *Bank Accessibility*

Motivation

Only 9.3 percent of rural respondents in our dataset live in a community with a bank. Demirgüç-Kunt et al. (2018) argue that, worldwide, distance from banks is a barrier to access to finance for many. According to the 2017 Global Findex Survey, 22 percent of adults without a bank account said that a reason for not having an account was that financial institutions were too far away. Furthermore, Guiso et al. (2004) see the presence of banks in a community as an indicator of financial development. People who live in less developed communities with few bank branches are more likely to rely on cash for their daily transactions. In support of these arguments, we observe that the proportion of respondents who only hold cash as a safe asset (cash ratio) is 68.8 percent (74.1 percent) in communities without banks, but only 56.1 percent (62.2 percent) in communities with banks. Hence, if a health shock occurs, we expect people living in communities without a bank to hold more cash in anticipation of future medical expenses.

Testing

To test whether lack of bank accessibility drives the association between health shocks and cash holdings, we estimate separate models for respondents who live in communities with and without banks. We expect the association between health shocks and cash holdings to be stronger for the latter.

Main Results

Panel C of Table 7 presents the coefficients associated with *Acute* for respondents who live in communities with and without banks. All estimates are carried out using a fixed-effects estimator, weighted using the weights obtained from the entropy balancing approach. The coefficients suggest that for residents who live in communities without banks, the occurrence of an acute health shock is associated with a 4.9 pp higher likelihood of holding only cash as a safe asset (column 1), and with a 3.4 pp higher cash ratio (column 2). By contrast, the coefficient associated with *Acute* is not statistically significant for respondents living in communities with banks. These findings are in line with our expectations.

Additional Results

The CHARLS provides community-level information on bank presence only in the 2011 wave. Thus, the information about bank branches is taken from the 2011 wave of the survey and used forward in the 2013, 2015, and 2018 waves. This assumes that bank presence remained constant in all waves of the survey, which can be justified by noting that, according to data from the Central Bank of China, compared with 2011, the average number of bank branches per 1000 people only increased by 0.005, 0.012, and 0.013 units, respectively, in 2013, 2015, and 2018 (see Appendix Figure S2). However, we hereafter verify that our main results hold when using a time-varying city-level measure of bank accessibility.

To this end, we use city-level administrative data on bank branches taken from the China Stock Market and Accounting Research (CSMAR) database. This dataset includes the date in which each branch was established, the full address of each branch, as well as an indication of whether the branch is located in a rural area. We then calculate the number of rural bank branches per 1000 rural people for each city in each year.[10] Next, we match this city-level measure of bank presence to our CHARLS sample and create dummy variables to identify respondents with a relatively high and low level of bank accessibility. Specifically, a respondent is defined as characterized by a low (high) bank accessibility in a given year if he/she lives in a city that falls in the bottom (top) half of the distribution of rural bank branches per 1000 rural people in that year.

[10] It is noteworthy that the cities in our dataset are prefecture-level cities. These are administrative units typically made up of a main central urban area (i.e., the core city which typically has the same name as the prefectural-level city) surrounded by rural areas. Rural population at the city level is calculated by multiplying the city's total population by one minus the urbanization rate in the province (assuming the rate is the same for all cities within the province). Data on the city-level population and the provincial-level urbanization rate are obtained from the China Stock Market & Accounting Research (CSMAR) database.

Finally, we re-estimate Equation (1) separately for respondents living in cities characterized by high and low bank accessibility. The results are presented in Panel D of Table 7.[11] The coefficients suggest that, in line with our previous findings, for residents who live in communities with low bank accessibility, the occurrence of an acute health shock is associated with a 7.2 pp higher likelihood of holding only cash as a safe asset (column 1), and with a 4.8 pp higher cash ratio (column 2). By contrast, the coefficient associated with *Acute* is never statistically significant in the sub-sample of respondents living in cities with high bank accessibility.

9. Conclusion

Using the 2013–2018 waves of the CHARLS, we highlight a higher probability of holding only cash as a safe asset and a higher cash ratio for those middle-aged and elderly respondents who experienced the onset of an acute health condition within the last 2 years in rural China. Furthermore, the effects of health shocks on cash holdings that we identify are only apparent for those respondents who do not benefit from immediate reimbursement of their healthcare expenses and for respondents who live in communities without banks. This suggests that ex-post reimbursement of medical expenses and lack of bank accessibility may drive the association we documented between health shocks and cash holdings.

Our paper has three main strengths. First, we are the first to explore the association between health shocks and portfolio choice focusing on cash holdings. We believe it is particularly important to understand the determinants of cash holdings in the Chinese context, considering that, by 2017, 225 million Chinese adults did not have a bank account, meaning that they mainly held cash in their portfolios (Demirgüç-Kunt et al., 2018). Moreover, our findings could be relevant to other countries such as India, Pakistan, Indonesia, Nigeria, Mexico, and Bangladesh, which, together with China, are home to nearly half the world's unbanked population (Demirgüç-Kunt et al., 2018). Second, we make use of a wide range of econometric techniques within a panel data setting. Uniquely, we account not only for correlated individual effects in several specifications, but we also produce unbiased estimates of the effect of experiencing a health shock on cash holdings by making use of an events study setting, as well as by matching treatment and control observations that are similar in terms of observed characteristics. Third, rather than just highlighting the links between health shocks and cash holdings, we also uniquely investigate possible drivers of these links, focusing on the role of the type of reimbursement of healthcare expenses that patients receive from their insurance scheme and the role of bank accessibility.

Our study has a few potential limitations. First, the data used to identify the presence of a health shock are self-reported survey data, which could be affected by measurement error. Yet, Gupta et al. (2015) show that individuals are not likely to misreport the presence or new diagnosis of specific conditions. Second, our analysis

[11] The number of observations in Panel D of Table 7 is slightly smaller than in the other Panels of the same Table due to missing values in the number of bank branches per 1000 people in some cities.

is based on a relatively short time horizon. However, Smith (2005) shows that most of the adjustment to acute health shocks takes place immediately. Third, our dataset only includes respondents aged 45 and above, missing therefore the dynamics for younger generations. Yet, older respondents' health is more likely to deteriorate, making the analysis of the effects of health shocks more relevant for this group of the population. Finally, due to data limitations, our paper does not take into account the fact that, in recent years, the use of mobile money through digital platforms, such as Alipay and WeChat, has become widespread in China. According to Demirgüç-Kunt et al. (2018), digital financial services might shrink the distance between financial institutions and their customers. As a result, they are likely to facilitate consumption smoothing in the presence of shocks and to reduce the need to hold cash (Lai et al., 2020). Yet, in the Chinese context, only people with bank accounts can access these platforms as providing bank details is a necessary step for completing the real name authentication needed to open WeChat/Alipay accounts (Huang et al., 2020; Shy, 2023). Hence, although our findings relative to the cash ratio may be affected by our failure to take mobile money into account, our results on the probability of holding only cash as a safe asset are not.

As money held in the form of cash is not deposited in financial institutions and does not circulate in the economy, a high reliance on savings in the form of cash has the potential for adverse effects on economic development (Levine & Zervos, 1998; Stix, 2013). Furthermore, too much cash in an economy has been associated with crime, tax evasion, and activities in the underground economy (Lahiri, 2020). The government should therefore devise policies aimed at reducing cash holdings in rural China. Based on our findings, these policies could operate by efficiently protecting households from the financial consequences of health shocks through a reduction of out-of-pocket healthcare costs or by providing immediate reimbursement of such costs more frequently. Steps in this direction have already been taken with the consolidation of the NCMS and URBMI into a new scheme called the Urban and Rural Resident Basic Medical Insurance scheme (URRBMI), which was initiated in 2016 and fully implemented in 2021 (Ren et al., 2022). The consolidation was aimed at gradually reducing the urban–rural gap and ensuring equitable benefits for urban and rural residents (State Council, 2016). As the Chinese rural population is likely to take time to understand the new consolidated insurance program, educating them so as to make them fully aware of the new benefits available to them (e.g., higher/quicker reimbursement of their medical expenditures) may be helpful in reducing their need to hold cash.

Policies aimed at enhancing bank presence in communities could also be helpful in reducing rural Chinese residents' cash holdings, as would policies aimed at enhancing rural residents' financial literacy. These would encourage people to become more reliant on banks, as well as on card and mobile payments, thus reducing their need to hold cash.

REFERENCES

Allen, F., Qian, J., & Gu, X. (2017). An overview of China's financial system. *Annual Review of Financial Economics*, *9*, 191–231.

Ameriks, J., & Zeldes, S. P. (2000). *How do household portfolio shares vary with age?* TIAA-CREF Working Paper. TIAA-CREF Institute.
Anderson, K. H., & Burkhauser, R. V. (1985). The retirement-health nexus: A new measure of an old puzzle. *Journal of Human Resources*, *20*, 315–30.
Angrisani, M., Atella, V., & Brunetti, M. (2018). Public health insurance and household portfolio choices: Unravelling financial "side effects" of medicare. *Journal of Banking and Finance*, *93*, 198–212.
Apeti, A. E., & Edoh, E. D. (2023). Tax revenue and mobile money in developing countries. *Journal of Development Economics*, *161*, 103014.
Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*, 29–51.
Atella, V., Brunetti, M., & Maestas, N. (2012). Household portfolio choices, health status and health care systems: A cross-country analysis based on SHARE. *Journal of Banking and Finance*, *36*, 1320–35.
Banks, J., O'Dea, C., & Oldfield, Z. (2010). Cognitive function, numeracy and retirement saving trajectories. *Economic Journal*, *120*, F381–F410.
Beck, T., Levine, R., & Loayza, N. (2000). Finance and the sources of growth. *Journal of Financial Economics*, *58*, 261–300.
Berkowitz, M. K., & Qiu, J. (2006). A further look at household portfolio choice and health status. *Journal of Banking and Finance*, *30*, 1201–17.
Bertocchi, G., Brunetti, M., & Torricelli, C. (2011). Marriage and other risky assets: A portfolio approach. *Journal of Banking and Finance*, *35*, 2902–15.
Black, N., De Gruyter, E., Petrie, D., & Smith, S. (2021). Altruism born of suffering? The impact of an adverse health shock on pro-social behaviour. *Journal of Economic Behavior and Organization*, *191*, 902–15.
Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*, 115–43.
Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, *115*, 707–47.
Calvet, L. E., Campbell, J. Y., & Sodini, P. (2009). Measuring the financial sophistication of households. *American Economic Review*, *99*, 393–8.
Calvet, L. E., & Sodini, P. (2014). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *Journal of Finance*, *69*, 867–906.
Cardak, B. A., & Wilkins, R. (2009). The determinants of household risky asset holdings: Australian evidence on background risk and other factors. *Journal of Banking and Finance*, *33*, 850–60.
Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior and Organization*, *83*, 50–8.
Chen, S., Ma, H., Teng, H., & Wu, Q. (2022). Banking liberalization and corporate tax planning: Evidence from natural experiments. *Journal of Corporate Finance*, *76*, 102264.
Cheng, T. C., Li, J., & Vaithianathan, R. (2019). Monthly spending dynamics of the elderly following a health shock: Evidence from Singapore. *Health Economics*, *28*, 23–43.
Christelis, D., Jappelli, T., & Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review*, *54*, 18–38.
Christiansen, C., Joensen, J. S., & Rangvid, J. (2015). Understanding the effects of marriage and divorce on financial investments: The role of background risk sharing. *Economic Inquiry*, *53*, 431–47.
Clark, A., & Etilé, F. (2002). Do health changes affect smoking? Evidence from British panel data. *Journal of Health Economics*, *21*, 533–62.
Cocco, J. F., Gomes, F. J., & Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *Review of Financial Studies*, *18*, 491–533.
Coile, C., & Milligan, K. (2009). How household portfolios evolve after retirement: The effect of aging and health shocks. *Review of Income and Wealth*, *55*, 226–48.
Cong, W., Gao, H., Ponticelli, J., & Yang, X. (2019). Credit allocation under economic stimulus: Evidence from China. *Review of Financial Studies*, *32*, 3412–60.
Correira, S. (2015). Singletons, cluster-robust standard errors and fixed effects: a bad mix. Technical Note, Duke University. http://scorreia.com/research/singletons.pdf
De Nardi, M., French, E., & Jones, J. B. (2010). Why do the elderly save? The role of medical expenses. *Journal of Political Economy*, *118*, 39–75.
Decker, S., & Schmitz, H. (2016). Health shocks and risk aversion. *Journal of Health Economics*, *50*, 156–70.
Dehejia, R., & Wahba, S. (2002). Propensity score matching methods for non-experimental causal studies. *Review of Economics and Statistics*, *84*, 151–61.
Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2018). *The global findex database 2017: Measuring financial inclusion and the fintech revolution*. World Bank.

Dobkin, C., Finkelstein, A., Kluender, R., & Notowidigdo, M. J. (2018). The economic consequences of hospital admissions. *American Economic Review*, *108*, 308–52.
Døskeland, T., & Kvaerner, J. S. (2022). Cancer and portfolio choice: Evidence from Norwegian register data. *Review of Finance*, *26*, 407–42.
Egger, P., Strecker, N. M., & Zoller-Rydzek, B. (2020). Estimating bargaining-related tax advantages of multinational firms. *Journal of International Economics*, *122*, 103258.
Feng, X., Lu, B., Song, X., & Ma, S. (2019). Financial literacy and household finances: A Bayesian two-part latent variable modeling approach. *Journal of Empirical Finance*, *51*, 119–37.
Fiebig, D., van Gool, K., Hall, J., & Mu, C. (2021). Health care use in response to health shocks: Does socio-economic status matter? *Health Economics*, *30*, 3032–50.
Fungáčová, Z., & Weill, L. (2018). Trusting banks in China. *Pacific Economic Review*, *23*, 131–49.
García-Gómez, P. (2011). Institutions, health shocks and labour market outcomes across Europe. *Journal of Health Economics*, *30*, 200–13.
Glaeser, E., Huang, W., Ma, Y., & Shleifer, A. (2017). A real estate boom with Chinese characteristics. *Journal of Economic Perspectives*, *31*, 93–116.
Goldman, D., & Maestas, N. (2013). Medical expenditure risk and household portfolio choice. *Journal of Applied Econometrics*, *28*, 527–50.
Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, *31*, 93–118.
Guiso, L., & Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, *6*, 1109–50.
Guiso, L., Sapienza, P., & Zingales, L. (2004). Does local financial development matter? *Quarterly Journal of Economics*, *119*, 929–69.
Gupta, N. D., Kleinjans, K. J., & Larsen, M. (2015). The effect of a severe health shock on work behavior: Evidence from different health care regimes. *Social Science and Medicine*, *136*, 44–51.
Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, *20*, 25–46.
Halko, M. L., Kaustia, M., & Alanko, E. (2012). The gender effect in risky asset holdings. *Journal of Economic Behavior and Organization*, *83*, 66–81.
Heaton, J., & Lucas, D. (2000). Portfolio choice in the presence of background risk. *Economic Journal*, *110*, 1–26.
Hong, E., & Hanna, S. D. (2014). Financial planning horizon: A measure of time preference or a situational factor? *Journal of Financial Counseling and Planning*, *25*, 184–96.
Huang, Y., Wang, X., & Wang, X. (2020). Mobile payment in China: Practice and its effects. *Asian Economic Papers*, *19*, 1–18.
Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*, 5–86.
Ivković, Z., Sialm, C., & Weisbenner, S. (2008). Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis*, *43*, 613–55.
Jones, A. M., Rice, N., & Zantomio, F. (2020). Acute health shocks and labour market outcomes: Evidence from the post-crash era. *Economics and Human Biology*, *36*, 100811.
Karlan, D., Ratan, A. L., & Zinman, J. (2014). Savings by and for the poor: A research review and agenda. *Review of Income and Wealth*, *60*, 36–78.
Lahiri, A. (2020). The great indian demonetization. *Journal of Economic Perspectives*, *34*, 55–74.
Lai, J., Yan, I. K. M., Yi, X., & Zhang, H. (2020). Digital financial inclusion and consumption smoothing in China. *China and the World Economy*, *28*, 64–93.
Lai, S., Shen, C., Xu, Y., Yang, X., Si, Y., Gao, J., Zhou, Z., & Chen, G. (2018). The distribution of benefits under China's new rural cooperative medical system: Evidence from western rural China. *International Journal for Equity in Health*, *17*, 137.
Lee, D. (2018). Effects of health insurance coverage on household financial portfolio: Evidence from the affordable care act. *Economics Letters*, *171*, 176–9.
Lei, X., & Lin, W. (2009). The new cooperative medical scheme in rural China: Does more coverage mean more service and better health? *Health Economics*, *18*, S25–S46.
Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. EconPapers. https://econpapers.repec.org/software/bocbocode/S432001.htm
Levine, R., & Zervos, S. (1998). Stock markets, banks, and economic growth. *American Economic Review*, *88*, 537–58.
Li, J., Li, Q., & Wu, X. (2020). Financial literacy, household portfolio choice, and investment return. *Pacific-Basin Finance Journal*, *62*, 101370.
Liu, G. G., Vortherms, S. A., & Hong, X. (2017). China's health reform update. *Annual Review of Public Health*, *38*, 431–48.

Ma, X., & Cen, Y. (2017). Public health insurance system reform and its impact on health service utilization in rural China: Evidence from CHNS 2000 and 2011. *Chinese Studies*, *6*, 85–107.

Meng, Q., Fang, H., Liu, X., Yuan, B., & Jin, X. (2015). Consolidating the social health insurance schemes in China: Towards an equitable and efficient health system. *Lancet*, *386*, 1484–92.

Miller, S., Johnson, N., & Wherry, L. R. (2021). Medicaid and mortality: New evidence from linked survey and administrative data. *Quarterly Journal of Economics*, *13*, 1783–829.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, *46*, 69–85.

National Health and Family Planning Commission, P.R. China. (2014). *China health and family planning yearbook*. Peking Union Medical College Press [In Chinese].

Papke, L. E., & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, *145*, 121–33.

Peng, R., Zhao, M., & Wang, L. (2014). Financial inclusion in the People's Republic of China: Achievements and challenges. *Financial Inclusion in Asia*, *7*, 7–43.

Peress, J. (2004). Wealth, information acquisition, and portfolio choice. *Review of Financial Studies*, *17*, 879–914.

Prina, S. (2015). Banking the poor via savings accounts: Evidence from a field experiment. *Journal of Development Economics*, *115*, 16–31.

Ren, Y., Zhou, Z., Cao, D., Ma, B. H., Shen, C., Lai, S., & Chen, G. (2022). Did the integrated urban and rural resident basic medical insurance improve benefit equity in China? *Value in Health*, *25*, 1548–58.

Riekhoff, A. J., & Vaalavuo, M. (2021). Health shocks and couples' labor market participation: A turning point or stuck in the trajectory? *Social Science and Medicine*, *276*, 113843.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, *9*, 86–136.

Rosen, H. S., & Wu, S. (2004). Portfolio choice and health status. *Journal of Financial Economics*, *72*, 457–84.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, *39*, 33–8.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*, 169–88.

Shoji, M., Cato, S., Ito, A., Iida, T., Ishida, K., Katsumata, H., & Mori McElwain, K. (2022). Mobile health technology as a solution to self-control problems: The behavioral impact of COVID19 contact tracing apps in Japan. *Social Science and Medicine*, *306*, 115142.

Shy, O. (2023). Cash is alive: How economists explain holding and use of cash. *Journal of Economic Literature*, *61*, 1465–520.

Sicular, T., Ximing, Y., Gustafsson, B., & Shi, L. (2007). The urban–rural income gap and inequality in China. *Review of Income and Wealth*, *53*, 93–126.

Smith, J. P. (2005). Consequences and predictors of new health events. In *Analyses in the economics of ageing* (pp. 213–40). University of Chicago Press.

State Council. (2016). State council advisory on consolidating urban, rural resident basic medical insurance. [In Chinese]. http://www.gov.cn/zhengce/content/2016-01/12/content_10582.htm

Stix, H. (2013). Why do people save in cash? Distrust, memories of banking crises, weak institutions and dollarization. *Journal of Banking and Finance*, *37*, 4087–106.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *25*, 1–21.

Sui, Y., & Niu, G. (2018). The urban-rural gap of Chinese household finance. *Emerging Markets, Finance and Trade*, *54*, 377–92.

Sundmacher, L. (2012). The effect of health shocks on smoking and obesity. *European Journal of Health Economics*, *13*, 451–60.

Verbeek, M., & Nijman, T. E. (1992). Testing for selectivity bias in panel data models. *International Economic Review*, *33*, 681–703.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *Journal of Finance*, *56*, 433–70.

Wang, Q., Zhang, D., & Hou, Z. (2016). Insurance coverage and socioeconomic differences in patient choice between private and public health care providers in China. *Social Science and Medicine*, *170*, 124–32.

Wang, Z., Li, X., Chen, M., & Si, L. (2018). Social health insurance, healthcare utilization, and costs in middle-aged and elderly community-dwelling adults in China. *International Journal for Equity in Health*, *17*, 17.

Watson, S. K., & Elliot, M. (2016). Entropy balancing: A maximum-entropy reweighting scheme to adjust for coverage error. *Quality and Quantity*, *50*, 1781–97.

Wooldridge, J. M. (2019). Correlated random effects models with unbalanced panels. *Journal of Econometrics*, *211*, 137–50.
World Health Organization. (2021). Ageing and health. https://www.who.int/china/health-topics/ageing
Wu, S. (2003). The effects of health events on the economic status of married couples. *Journal of Human Resources*, *38*, 219–30.
Yin, X., Abruquah, L., & Ding, Y. (2019). Dynamics of life satisfaction among rural elderly in China: The role of health insurance policies and intergenerational relationships. *Sustainability*, *11*, 701–20.
Zhang, A., Nikoloski, Z., & Mossialos, E. (2017). Does health insurance reduce out-of-pocket expenditure? Heterogeneity among China's middle-aged and elderly. *Social Science and Medicine*, *190*, 11–9.
Zhang, J., Gan, L., Xu, L. C., & Yao, Y. (2014). Health shocks, community elections, and household income: Evidence from rural China. *China Economic Review*, *30*, 155–68.
Zhang, X., & Zhang, L. (2022). The impact of instant reimbursement of cross-regional medical services on hospitalization costs incurred by the floating population–evidence from China. *Healthcare*, *10*, 1099.
Zhong, H. (2011). Effect of patient reimbursement method on health-care utilization: Evidence from China. *Health Economics*, *20*, 1312–29.
Zhou, Q., Liu, G. G., Sun, Y., & Vortherms, S. A. (2016). The impact of health insurance cost-sharing method on healthcare utilization in China. *China Journal of Social Work*, *9*, 38–61.
Zou, J., & Deng, X. (2019). Financial literacy, housing value and household financial market participation: Evidence from urban China. *China Economic Review*, *55*, 52–66.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Data SI. Replication kit.
Appendix SI. Additional figures and tables.
- **Figure S1.** Entropy balancing weights.
- **Figure S2.** Number of bank branches per 1,000 people (2008–2018).
- **Table S1.** Summary statistics.
- **Table S2.** First-stage regression for PSM model.
- **Table S3.** Balancing tests for the validity of the PSM approach.
- **Table S4.** Baseline specifications estimated on a matched sample.
- **Table S5.** Summary statistics of relevant variables before and after entropy balancing weighting.