World Development 153 (2022) 105838

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Income shock and food insecurity prediction Vietnam under the pandemic

Khoa Vu [a,*], Nguyen Dinh Tuan Vuong [b], Tu-Anh Vu-Thanh [c], Anh Ngoc Nguyen [d]

[a] Department of Applied Economics, University of Minnesota, United States
[b] Department of Agricultural and Applied Economics, University of Wisconsin-Madison, United States
[c] Fulbright School of Public Policy and Management, Fulbright University Vietnam, Vietnam
[d] Development and Policies Research Center (DEPOCEN), Vietnam

ARTICLE INFO

Article history:
Accepted 28 January 2022

Keywords:
Food insecurity
Income shock
COVID-19
Vietnam

ABSTRACT

As COVID-19 threatens the food security of vulnerable populations across the globe, there is an increasing need to identify places that are affected most in order to target aid. We propose a two-step approach to predict changes in food insecurity risk caused by income shocks at a granular level using existing household-level data and external information on aggregate income shocks. We apply this approach to assess changes in food insecurity risk during the pandemic in Vietnam. Using national household survey data between 2010 and 2018, we first estimate that a 10% decrease in income leads to a 3.5% increase in food insecurity. We then use the 2019 national Labor Force Survey to predict changes in the share of food-insecure households caused by the income shocks during the pandemic for 702 districts. We find that the small, predicted change in food insecurity risk at the national level masks substantial variation at the district level, and changes in food insecurity risk are larger among young children. Food relief policies, therefore, should prioritize a small number of districts predicted to be severely affected.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Food insecurity has become a major economic consequence of the global pandemic as lockdown and social distancing are widely adopted to contain the COVID-19 virus.[1] During a lockdown, many workers lose part of their incomes or their jobs as they withdraw from social interaction, which, in turn, severely affect their ability to afford food (Devereux et al., 2020). Although several COVID-19 vaccines were successfully developed and approved in early 2021, the world continues to experience new waves of infection with newer and more transmissible variants, especially among developing countries that lack the resources to purchase and distribute vaccines to their citizens. Thus, food insecurity will continue to be an issue until lockdown orders and/or withdrawals from work and social interaction are no longer necessary, i.e. when herd immunity is achieved through vaccine-induced or natural immunity.

Given the severity of the economic impacts of the pandemic, governments and international organizations have pursued various actions to provide financial support and food aid to vulnerable populations (Gentilini, Almenfi, Orton, & and Dale, 2020), which raises an important question: how should governments and organizations effectively target places or households that need their support most within a country? It is now clear that the economic impacts of COVID-19 vary across populations, geography, and economic sectors (The World Bank, 2020; Fund, 2020). Prioritizing places affected more by the pandemic will allow governments or international organizations to provide more support to people needing them most instead of extending resources to people who were relatively unaffected. This is especially true for low- and middle-income countries with limited resources.

A common approach is to send aid to households with a pre-pandemic poverty or low-income status.[2] While it is simple to target poor households, not all poor households experience food insecurity during the pandemic. For example, households whose

* Corresponding author.
E-mail addresses: vuxxx121@umn.edu (K. Vu), nguyen.vuong@wisc.edu (N.D.T. Vuong), anh.vu@fulbright.edu.vn (T.-A. Vu-Thanh), ngocanh@depocen.org (A.N. Nguyen).

[1] See Devereux, Béné, and Hoddinott (2020); Brown, Ravallion, and and Van De Walle, 2020; Mishra and Rampal, 2020; Ahn and Norwood, 2020; Paslakis, Dimitropoulos, and Katzman, 2020; Reardon, Bellemare, and Zilberman, 2020; Fund, 2020.

[2] The goal of this targeting approach can be to support the most vulnerable population or to compensate for their income losses.

https://doi.org/10.1016/j.worlddev.2022.105838
0305-750X/© 2022 Elsevier Ltd. All rights reserved.

members work in unaffected sectors are not suffering, while near-poor households may still suffer a food insecurity shock if their members work in a severely affected sector.[3] This issue is particularly severe in developing countries where data on poverty status may be outdated and data on income does not include informal economic activities (Aiken, Bellue, Karlan, Udry, & Blumenstock, 2021). A better method is to combine household targeting with geographic targeting: government (or international organizations) can first identify and allocate an appropriate budget and resources to localities that are affected most, so that local governments or field agents from international organizations can identify and distribute aid to households or individuals that require assistance. This practice is used by the United Nations World Food Programme (World Food Programme, 2015).

Identifying which places are affected more, however, would require conducting a large-scale survey to determine the percentage of food-insecure households in different areas across a country, which can be both costly and time-consuming. Facing this constraint, most existing studies only document food insecurity shocks during the pandemic at the national or regional level.[4] Several innovative studies propose predicting and targeting food insecurity at a more granular level using remote sensing data on geography and weather (e.g., Andree, Chamorro, Kraay, Spencer, & and Wang, 2020; von Carnap, 2021; Zhou, Lentz, Michelson, Kim, & and Baylis, 2021) or combining such data with phone surveys and mobile phone usage (Blumenstock, Cadamuro, & On, 2015; Aiken et al., 2021). These studies tap into a significant amount of data that are traditionally underutilized, but they also require using cutting-edged machine learning technique to process satellite photos and make prediction based on a large number of factors. The complex nature of these advanced methods, however, can pose a significant barrier for policymakers to adopt widely. These methods rely heavily on the expertise of the researchers and modelers to make decisions on which data and methods to use as well as decisions about variables, parameters, and assumptions specific to each method.[5]

We propose an alternative approach that also predicts and target food insecurity at the granular level and only relies on a common and widely-used method in the econometric toolkit. The first step is to estimate the causal effect of household income on food insecurity. The standard OLS estimation may suffer potential bias as food insecurity can simultaneously affect income by lowering productivity, so we use local exposure to national employment changes to instrument for household income. The second step is to combine this estimate and external information on aggregate income shocks at the sector level during the pandemic to predict changes in food insecurity at the locality level. Because we only focus on predicting changes, not levels of, food insecurity, we only require information about how income changes at the sector level. Under the assumption that the pandemic mainly affects food insecurity through the income loss channel, this method can provide a reasonable and timely prediction that policymakers can quickly act on.

We apply this approach to predict and assess potential food insecurity shocks during the pandemic in Vietnam. In 2020, Vietnam was among the most successful countries at containing new infections with strict measures of mass testing, targeted lockdowns, and quarantine policies. However, in 2021, the country entered a new and more severe wave of infection, forcing more extensive lockdowns in several parts of the country. Although food insecurity is reported as one of the major concerns of households during this period (Yang, Panagoulias, & and Demarchi, 2020), little is known about which places in Vietnam are affected more in terms of food insecurity, preventing the government and international organizations from targeting and distributing aid. We address this problem by predicting food insecurity shocks due to changes in income for 702 districts in Vietnam, allowing identification of districts that are likely affected most.

We first use the 2010–2018 Vietnam Household Living Standard Survey (VHLSS) to estimate the effect of income shock on food insecurity. We then use information from the World Bank on pandemic income shocks and the 2019 Labor Force Survey (LFS) to predict changes in food insecurity risk for 702 districts of Vietnam. The share of food-insecure households is predicted to rise by 0.82 percentage points, but a small number of districts are predicted to experience increases as large as 7.86 percentage points. We also predict an increase of 0.997 percentage points in the share of food-insecure children under 5, and a few districts are predicted to experience an increase as large as 19.33 percentage points. These predictions suggest that the average impact of the income shock during the pandemic in Vietnam may be small, but certain parts of the country are likely affected more severely than others.

Our study makes two contributions to the literature. First, it contributes to a growing number of studies on predicting food insecurity at the locality level for targeting purposes, which are part of a broader literature on interventions and disaster-relief policies to address food insecurity during emergencies.[6,7] Whether an intervention is effective depends crucially on the quality of the targeting approach (Barrett, 2010; Lentz & Barrett, 2013): targeting the wrong places is costly, while missing the food-insecure places can have long-term negative consequences for food-insecure households (Brown, Ravallion, & van de Walle, 2018). Although the literature offers several machine learning approaches to make timely and granular prediction on food insecurity for targeting purposes during crises that are unrelated to COVID-19 (e.g., Lentz, Michelson, Baylis, & Zhou, 2019; Andree et al., 2020; Zhou et al., 2021), only two studies focus on developing methods specific to the pandemic. Gundersen, Hake, Dewey, and Engelhard (2020) employ an OLS-based approach to predict changes in food insecurity at the county level in the US. Relatedly, Aiken et al. (2021) propose a new framework to combine data from satellite images and mobile phone networks with phone survey data to predict levels of poverty in Togo. Like Gundersen et al. (2020), our study also predict changes, not levels, of food insecurity. We extend the methods in Gundersen et al. (2020) to address the simultaneity bias in the standard OLS-based approach. Our proposed method is more suitable for developing countries than that of Gundersen et al. (2020) because it employs datasets that are typically available in these countries, including population censuses, national household surveys, and labor force surveys.

Second, our study can help policymakers address the rising concerns about food insecurity in Vietnam and other countries.[8] In the

[3] Gundersen, Kreider, and Pepper (2011) show that in the US, 65% of households close to the poverty line are food secure, while 10%–20% of households with income twice as high as the poverty line are food insecure.

[4] See, e.g., Ahn and Norwood, 2020; Wolfson and Leung, 2020; Bitler, Hoynes, and and Schanzenbach, 2020; Aggarwal et al., 2020; Amare, Abay, Tiberti, and Chamberlin, 2021; Hirvonen, de Brauw, and Abate, 2021; Gupta, Zhu, Doan, Michuda, and Majumder, 2021; Nguyen et al., 2021; Barrett, Upton, Tennant, and and Florella, 2021; Gatto and Islam, 2021 among many others.

[5] Zhou et al. (2021) document several modeling choices that can have important consequences on the prediction of these models.

[6] Several studies find that timely interventions such as cash or food transfers can reduce food insecurity and mitigate any negative effects of food insecurity, such as Nikulkov, Barrett, Mude, and Wein (2016); Gelli et al., 2017; Hidrobo, Hoddinott, Kumar, and Olivier, 2018; Christian, Kandpal, Palaniswamy, and Rao, 2019; Savy et al., 2020 and others. Lentz and Barrett (2013) show that food assistance policies have the highest returns when targeting children under 2 years old.

[7] Our study is also related to an extensive literature on factors of food insecurity such as rainfall (Niles & Brown, 2017), agricultural income (Kuma, Dereje, Hirvonen, & Minten, 2019), farm production (Jones, Shrinivas, & Bezner-Kerr, 2014), and economic vulnerability (Chaaban, Ghattas, Irani, & Thomas, 2018).

[8] Because we only focus on food insecurity prediction, our study is only indirectly linked to an extensive literature on measuring the impacts of the pandemic on food insecurity especially in developing countries. See Picchioni, Goulao, and Roberfroid (2021) and Béné et al. (2021) for recent systematic reviews on this topic.

case of Vietnam, a national phone survey conducted by the World Bank indicates that 33% of households were concerned about not having enough food (Yang et al., 2020), yet which districts are affected more remains unknown. While Vietnam continues to combat new waves of infection with extensive lockdown orders, it becomes increasingly important for policymakers to sketch a rapid and well-targeted response to alleviate the pandemic's economic impacts in Vietnam. Using information about income shocks in 2020, our study maps the projected food insecurity shocks at the district level, allowing policymakers to identify districts that may require more assistance than others. Policymakers can also apply our proposed method to predict food insecurity shocks as the pandemic continues or for other shocks in Vietnam and other developing countries.

The rest of this paper is organized as follows. Section 2 summarizes the impacts of COVID-19 on the Vietnamese economy. Section 3 proposes an empirical framework to predict changes in food insecurity due to sector-specific income shocks and describes the data used. Section 4 reports the regression results. Section 5 applies the framework to predict food insecurity changes in Vietnam in 2020 and discusses policy implications. Section 6 summarizes the lessons from our predictions and discusses different caveats of the proposed approach.

2. Impacts of COVID-19 on the Vietnamese economy

In this section, we briefly discuss how the pandemic evolved in Vietnam and its economic consequences from January 2020 to July 2021. The first five cases of COVID-19 were detected during the last week of January 2020. Since then, several public health measures have been taken to contain the virus's spread, which in turn has affected the economy. On February 1, 2020, flights between China and Vietnam were canceled (Tuoi Tre online, 2020). The borders between the two countries have also been tightened, disrupting agricultural exports from Vietnam to China (BBC, 2020). Moreover, this strongly affected tourism, an important industry in Vietnam. According to the Vietnam National Administration of Tourism (2021b), Vietnam received 18 million international tourists in 2019 and collected revenue of $27.5 billion in 2018, which is about 11% of 2018 GDP (Vietnam National Administration of Tourism, 2021a). Following the travel restrictions imposed in February 2020, the Vietnamese tourism industry lost 32% of its international customers (mainly from China). Starting in March 2020, all flights to Vietnam were canceled due to the outbreak in the EU and the US. The Vietnamese government also imposed a two-week social distancing order on April 1, 2020. These events created shocks on both demand and supply in the tourism sector as well as other related industries in the service sector.

The social distancing order in April also affected manufacturing industries by requiring additional distance between workers. According to the Ministry of Industry and Trade (MOIT), the industrial production index decreased by 18% in April 2020 and only recovered to pre-distancing levels in June 2020. Meanwhile, the outbreak also created a supply chain disruption in the manufacturing sector. According to a survey conducted by the Vietnamese General Statistics Office (GSO) in April 2020, 42.8% of surveyed firms reported a supply shortage, and this number increased to more than 70% in the garment industry (GSO, 2020b). This issue was partly resolved when China, Korea, and Japan reopened their economies, and Vietnam's imports in the first seven months of 2020 were lower than the same period in 2019, only by 2.9% (Ministry of Industry & Trade, 2020).

While the initial supply shock abated as the country successfully contained the first wave of the virus, Vietnam still faced lower demands from importing countries due to global supply chain issues. For instance, in the first seven months of 2020, the garment and footwear industries exported less than that of 2019 by 12%. To cope with lower demand due to the pandemic, firms reportedly cut costs by terminating labor contracts, laying off workers, cutting workloads, and/or cutting wages.[9] This in turn led to higher risks of food insecurity, especially for female and low-skilled workers.

The first wave in Vietnam ended in April 2020 when the country recorded no new local transmission cases for almost three months. The second wave, however, started in July 2020 when there were several new cases and deaths across multiple cities, forcing local lockdowns. Following the second wave, economic recovery slowed down. According to Bank (2020), the Vietnamese economy grew at a slower rate in August 2020 compared to previous months. Retail sales grew by 5.2% in July 2020 and only 2.3% in August 2020. FDI flow also decreased substantially; in July 2020, it was US$3.1 billion but dropped to US$720 million in August 2020. Within Vietnam, vulnerable workers continue to worry about future financial outcomes (Dang & Giang, 2020).

The third wave started in late January 2021 and quickly spreaded across the country, forcing several cities and provinces to lock down until March 2021 (Ministry of Health, 2021). In early May 2021, state and foreign media issued warnings about a potential fourth wave as new cases were detected in Hanoi, Vinh Phuc, and Ha Nam (BBC, 2021). A fourth wave of infection started in early May 2021, and Vietnam continued to struggle with containing it as of July 2021. This wave is the most severe so far, as the number of new cases from May accounts for 84% of total cases in Vietnam up to this point. The industrial powerhouses, including Ho Chi Minh City, Bac Ninh, and Bac Giang, were the epicenters of this wave. Targeted lockdowns and contract tracing were implemented across the country, affecting 15 million workers in the second and third quarters of 2021 (GSO, 2021). The full extent of the economic impacts of the fourth wave and the pandemic on Vietnam remains unclear as new variants such as the Omicron variant continue to emerge.

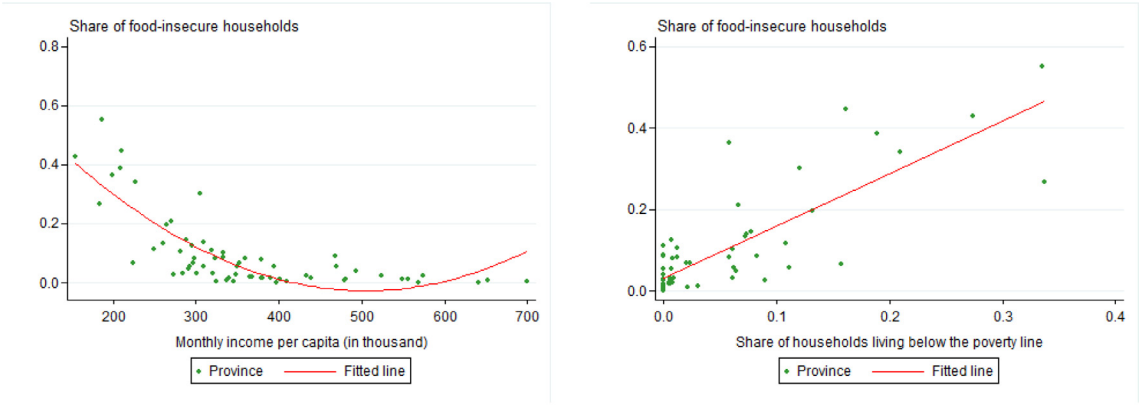
Given these economic impacts, one can expect the pandemic to have severe effects on food insecurity. As household income decreases due to the pandemic, households are less likely to afford nutritious food, which threatens their food security. We illustrate the relationship between food insecurity and income in Fig. 1 by plotting province-level average monthly income against shares of food-insecure households in (a) as well as province-level shares of poor households against shares of food-insecure households in (b). Provinces with higher average income and a lower share of poor households tend to have a lower share of food-insecure households. In the next section we discuss in detail how to predict food insecurity shocks due to the pandemic in Vietnam.

3. Empirical strategy

3.1. Data

As we will discuss in the next section, the method proposed in this study requires data for two steps: (1) estimating the causal relationship between household income and food insecurity and (2) predicting food insecurity changes at the district level by combining the estimates from the first step with external information about income shocks. We use the 2010–2018 Vietnam Household Living Standard Survey (VHLSS) for the estimation step. The VHLSS is a biannual survey conducted by the General Statistics Office of Vietnam (GSO). Each wave of data consists of nearly 9,400 house-

[9] In April 2020, the GSO reported that 66.8% of firms applied at least one of these measures GSO (2020b), and the LFS reports that 30.8 million labors were affected in quarter 2 (GSO, 2020a).



(a) Income versus food insecurity (b) Poverty versus food insecurity

Fig. 1. Correlation between province-level income and poverty with share of food-insecure households in 2018. Note: Figure (a) shows a scatter plot of the province-level share of food-insecure households and the average monthly income per capita. Figure (b) shows a scatter plot of the province-level share of food-insecure households and the share of poor households. A household is defined as food insecure when its share of calories from staple food exceeds 84% (see more in the next section). A household is defined as poor when its monthly income per capita falls below the national poverty line. The data are from the Vietnam Household Living Standard Survey in 2018.

hold observations and roughly 36,000 individual observations across the country. This dataset is well-suited to estimate the effect of household income on food insecurity status for two main reasons. First, it contains many variables related to household food consumption and food insecurity, allowing us to construct different measures of food insecurity based on nutritional quality.[10] Second, it covers all provinces over a long time period, which allows us to control for any long-term trends in income and food insecurity at the province level. However, there are two important caveats of the VHLSS that make it unsuitable for the prediction step: (1) the latest wave available is 2018, while the income shocks reported by the World Bank are a comparison between 2020 and 2019; and (2) the sampling of the study is not designed to calculate aggregates at the district level; that is, it does not have enough observations per district for aggregation.

Due to these limitations, we use the Labor Force Survey (LFS) in 2019 for the prediction step. The survey is conducted annually by the GSO and includes more than 800,000 individuals from about 200,000 households across 702 districts. Households are selected from the stratified random sampling method that is representative at the district level. The LFS focuses on the labor market information of individuals of legal working age, which includes employment status, income, and workload, and has demographic information on all household members. The LFS allows aggregating data at the district level, and the latest data available is for 2019, which is more comparable to 2020 than the latest VHLSS wave in 2018. However, the LFS data do not have information on food consumption, which is why we cannot use it for the estimation step. In Table 1, we provide summary statistics for household data from the 2010–2018 VHLSS and the 2016, 2018, and 2019 LFS.

Our instrumental variable is the share of adult employment at the district-industry level in 2009 times the national employment growth of each industry. We construct this variable using the 2009 Population and Housing Census. The census is conducted every ten years by the GSO and contains information of the whole population such as demographics, migration, educational attainment, employment, fertility, mortality, and housing quality. We use the 15% sample provided by Minnesota Population Center, 2019, which has about 14 million observations and is representative at the district level.

3.2. Measuring food insecurity

According to Bickel, Nord, Price, Hamilton, and and Cook, 2000, food security is when access to nutritionally adequate, safe, and socially acceptable foods is not limited or uncertain. Following this definition, there are three core concepts that define a food-insecure household: availability, accessibility, and utilization (Webb et al., 2006; Barrett, 2010). Depending on the data availability and the context, researchers can use proxy for these different aspects of food insecurity. In this study, we focus on measuring the utilization aspect of food insecurity, which reflects concerns about whether households make good use of the food to which they have access (Barrett, 2010).

Following Baylis, Fan, and Nogueira (2019) and Jensen and Miller (2010), we define a food-insecure household as a household with the share of calories from staples exceeding 84%. This approach follows Bennett's Law that people consume more nutrient-dense foods and reduce their consumption of calorie-dense staple foods as their income increases. As households prioritize their calorie requirement for basic activities, the shift from consuming calorie-dense staple food to more protein-dense non-staple food implies that households have achieved their desired calorie intake. This approach falls under the utilization concept of food insecurity, as it is concerned with whether households consume nutritionally essential or nutritionally inferior foods (Barrett, 2010). Jensen and Miller (2010) find that an adult in China would not meet his nutritional demand if the fraction of his calories from rice, wheat, and other staple food exceeds 84%. We also use this 84% of the staple calorie share to define a food-insecure household given that Vietnam and China are similar. With the conversion table from Thi, Simioni, and Thomas-Agnan (2018),[11] we calculate the monthly calorie intake by food categories and construct a dummy variable equal to zero when staple calorie share is below 84% and 1 when the share is above 84%. This binary variable is our main measure of food insecurity outcome.

There are other methods to measure food insecurity. One approach is using calorie intake to define food insecurity: researchers define the minimum requirement of calorie intake in a day. For

[10] In addition to containing a lot of information about household consumption, VHLSS is a rotating panel, where 50% of the sample appears in two consecutive surveys and a household can only be tracked for three waves. We take advantage of this unique feature to construct a subsample with household-level panel data for a household fixed effects specification.

[11] The table is a summary of National Institute of Nutrition (2007).

Table 1
Summary statistics.

Panel A: 2010–2018 VHLSS

Variables	2010		2012		2014		2016		2018	
	Mean	(S.E.)	Mean	(S.E.)	Mean	(S.E.)	Mean	(S.E.)	Mean	(S.E.)
Staple calorie share (%)	75.12	(11.40)	74.89	(10.84)	72.64	(11.60)	70.87	(11.16)	68.90	(11.76)
Constructed food insecurity (%)	21.83	(41.31)	19.78	(39.83)	14.12	(34.83)	11.04	(31.34)	7.81	(26.83)
Self-reported food insecurity (%)	4.64	(21.04)	3.74	(18.98)	2.62	(15.99)	1.85	(13.47)	1.23	(11.00)
Household dietary diversity score	10.01	(1.54)	10.07	(1.51)	10.16	(1.56)	10.25	(1.45)	10.26	(1.54)
Yearly income ('000,000, 2010 VND)	66.36	(117.25)	71.09	(78.90)	75.96	(76.34)	84.88	(83.76)	100.48	(95.18)
Household size	3.95	(1.56)	3.91	(1.56)	3.84	(1.57)	3.81	(1.60)	3.74	(1.60)
Urban (%)	28.26	(45.03)	28.80	(45.29)	29.76	(45.72)	30.33	(45.97)	30.30	(45.96)
No. observations	9,267		9,278		9,297		9,302		9,299	

Panel B: 2016, 2018 & 2019 LFS

Variables	2016		2018		2019	
	Mean	S.E.	Mean	S.E.	Mean	S.E.
Yearly income ('000,000, 2010 VND)	71.98	(72.17)	78.25	(71.78)	82.20	(71.84)
Household size	3.78	(1.55)	3.77	(1.57)	3.70	(1.56)
Urban	42.68	(49.46)	42.67	(49.46)	42.63	(49.45)
No. observations	206,385		208,905		212,040	

example, Abebaw, Admassie, Kassa, and Padoch (2020) define a household as food-insecure when it uses adult-equivalent 2,200 calories or less per day. However, such measures may not be appropriate because there are no consistent ways to calculate the adult-equivalent unit. Thi et al. (2018) compare the estimations on average adult-equivalent calorie intake at the household level from studies using different methods on 2004 VHLSS, and the estimation results can range between 2,300 to 3,300 calories per day across different methods. Given that the results are very sensitive to the choice of an adult-equivalent method, using calorie intake to measure food insecurity will lead to the same problem.

Another popular approach is using self-reported information on food insecurity. Within the sustainability dimension, researchers can measure food security by asking whether a household has had two meals per day in the last 12 months or if households have not had enough food to eat in any month within the last 12 months (e.g., Kuku, Gundersen, & Garasky, 2011; Ratcliffe, McKernan, & Zhang, 2011; Verpoorten, Arora, Stoop, & Swinnen, 2013; Gundersen, Kreider, & Pepper, 2017; Abebaw et al., 2020). In addition, this type of approach can measure the utilization dimension by asking households whether they consume less preferred foods or unbalanced meals (Ratcliffe et al., 2011).

The self-reported approach is problematic in our setting for two main reasons. First, a household only answers the question about food insecurity when it has been recognized as a "poor household" by the government in the last five years; in other words, a non-poor household is automatically not considered as food-insecure. This approach would underestimate the number of food-insecure households by excluding the near-poor group and lead to selection bias when estimating the effect of household income on food insecurity. If the government changes the definition of a poor household, it can also lead to artificial changes in the food insecurity prevalence.

Second, asking whether households have enough food to measure food insecurity can be misleading given the context of Vietnam. Being the second-largest rice exporter in the world, the price of rice in Vietnam can be as cheap as 22 cents per pound,[12] but nutrition-dense food is more expensive. For example, the unit price of pork is ten times more expensive than rice.[13] Therefore, nutritional deficiency is likely a more significant concern than the number of meals per day. Moreover, due to the short-term nature of lockdowns during the pandemic, households may fall back on staple food and reduce consumption of more expensive food to smooth food consumption (Hirvonen et al., 2021), which means composition of food consumption is likely more relevant than the overall consumption. Given these two reasons, we choose not to use the self-reported measure as the main measure of food insecurity.[14]

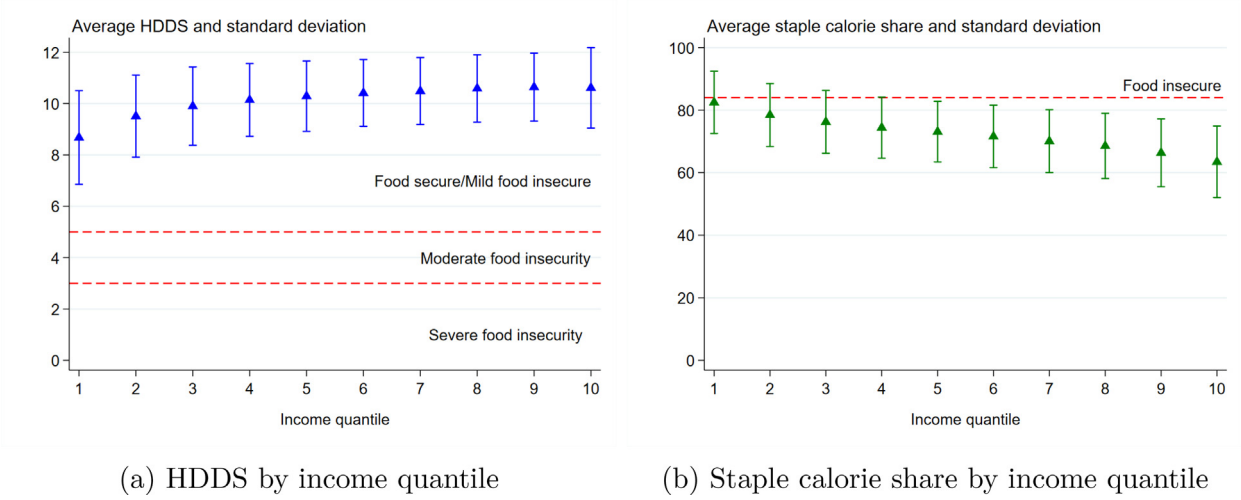
Another commonly used measure is the household dietary diversity score (HDDS) (Swindale & Bilinsky, 2016; Maxwell, Vaitla, & Coates, 2014). However, the HDDS only counts the number of food categories that households consume and does not account for the depth of food insecurity in terms of food quantity. In other words, the score would take into account households with high HDDSs that still suffer nutritional deficiency. HDDSs are also typically calculated using daily or weekly food consumption data; however we only have monthly data, so this might exaggerate the dietary diversity. Furthermore, dietary diversity could be explained by climate or culture; for example, tropical climates are associated with higher biodiversity and hence higher dietary diversity. For this reason, there is no universal HDDS threshold for a food-insecure household across different countries (Deitchler, Ballard, Swindale, & and Coates, 2010), as one threshold that works in one country may not work in another.

To illustrate the problem with the food insecurity threshold using the HDDS, in Fig. 2a we plot average HDDS scores by income group and the thresholds being used to determine food-insecure households by other food insecurity studies (e.g., Vaitla, Coates, & Maxwell, 2015; Vaitla et al., 2017; Lentz et al., 2019). The figure shows that even the lowest income group in our data has an average HDDS score that exceeds all of these thresholds. Therefore, these HDDS thresholds used in other studies or other countries cannot be applied in Vietnam's setting because all households would be classified as not food insecure.[15] Our main outcome uses the fraction of calories from staple food and the 84% cutoff from Jensen and Miller (2010); given Vietnam and China are relatively more similar, we do not observe such a problem in our approach, as seen in Fig. 2b.

[12] Source (in Vietnamese): Vietnambiz (2020b)
[13] Source (in Vietnamese): Vietnambiz (2020a)

[14] Another problem, although not as severe, is that self-reported data can be inconsistent over time. In the case of the VHLSS, a household is considered food-insure if the household does not have enough two meals a day at some points in 2012–2018, but in 2010 a household is food insecure if it reports that it does not have enough food.

[15] For this reason, we must pick a higher HDDS cutoff for food-insecure households for our analysis.



(a) HDDS by income quantile

(b) Staple calorie share by income quantile

Fig. 2. Food insecurity measures by income group, 2010–2018. Note: The figure shows the average and standard deviation for the Household Dietary Diversity Score (HDDS), and the staple calorie share is estimated for ten income quantiles. For the HDDS measure, we also show the common cutoffs to identify food-insecure households (e.g., Vaitla et al., 2015; Vaitla et al., 2017; Lentz et al., 2019). For the staple calorie share, we show the food-insecure cutoff estimated (Jensen & Miller, 2010) for Wuhan, China, which we also use in this study.

In order to check the robustness of our regression estimation, we also use the food insecurity measures from self-reporting and the dietary diversity score as alternative outcomes, both of which are binary variables. The self-reported food insecurity variable is equal to one if a household does not have enough two meals per day for one month or more, and zero otherwise. There is no universal cutoff for the HDDS to define a food-insecure household based on the dietary diversity score, but the official guideline of HDDS suggests that a meaningful target level of diversity can be set based on the average score of the richest 33% of households (Swindale & Bilinsky, 2006). We follow this guideline and set the cutoff value for HDDS at 10; that is, the HDDS-based food insecurity variable is one if the dietary diversity score is below 10, and zero if the score is above 10. This cutoff, however, likely results in an overestimated share of food-insecure households because this definition implies that only the richest 33% of households are not food-insecure.

To see how different food insecurity measures actually differ in our data, we use the 2010–2018 VHLSS to map the province-level share of food-insecure households over time using the (a) self-reported measure, (b) HDDS-based measure, and (c) staple calorie share-based measure in Fig. 3.For comparison, we also map each province's share of poor households, defined as those with a monthly income per capita lower than the national poverty line in Figure (d). We observe that provinces with a higher share of poverty also have a higher share of food-insecure households across all three measures. Provinces in remote and mountainous areas, including the Northeast, Northwest, and Central Highlands, tend to have higher poverty and food insecurity rates. There are stark differences between different measures of food insecurity. On one hand, the rates of self-reported food insecurity tend to be substantially lower than the poverty rates across provinces and years, confirming the selection bias problem that we discuss above. On the other hand, the rates of HDDS-based food insecurity tend to be substantially higher than the poverty rates, while the staple calorie share-based food insecurity rates track closely with the poverty rates. We also observe that poverty declines over time;[16] at the same time, we observe that the self-reported food insecurity rises (especially in 2016)[17] while the HDDS-based food insecurity stays relatively stable. In contrast, the changes of food insecurity based on the staple calorie share approach also track the movement of the poverty rates. This suggests that measuring food insecurity using the staple calorie share approach is valid in the context of Vietnam. We further discuss the validity of this measure in Appendix B.

3.3. Econometric model

We propose a simple approach to predict food insecurity risk caused by major income shocks using past household living standard surveys and external information about sector-specific income shocks, which is typically reported by the government or international organizations. Although we focus on income shocks during a pandemic, this method can also be applied to income shocks during other emergencies. Let Y_{it} denote the binary outcome variable that indicates whether household i is food insecure in period t, consider the following linear probability model:

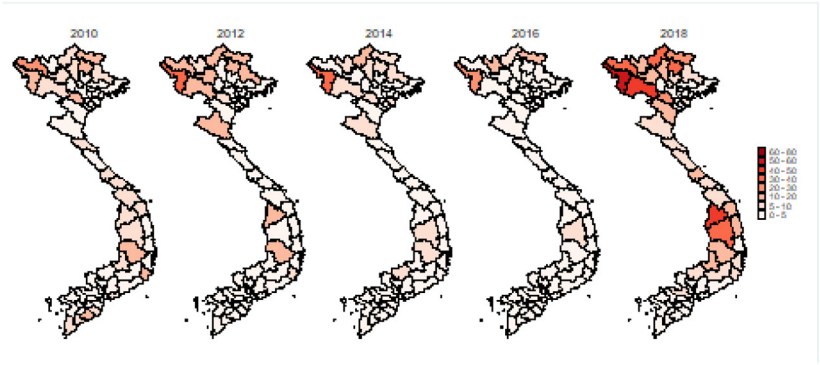
$$Pr(Y_{it} = 1|income_{it}, \mathbf{Z_{it}}) = \beta.income_{it} + \mathbf{Z_{it}}\theta, \quad (1)$$

where $Pr(Y_{it} = 1|income_{it}, \mathbf{Z_{it}})$ is the conditional probability of being food insecure; $income_{it}$ denotes the household's real income in period t; Z_{it} denotes other time-varying household-level factors driving food insecurity, and β represents the causal effect of the household's income on food insecurity. Assume that household i provides labor in sector s and receives a salary of e_{ist} in period t. Then household income is defined as the sum of the salary received in each sector s in which the household provides labor: $income_{is} = \sum_j e_{ist}$. To understand how a major income shock can affect food insecurity, we first assume that the pandemic changes the salary of an average worker in sector s by z_s. Let $income_{it}^{pre}$ and $income_{it}^{post}$ denote the household's income in the pre-shock and post-shock periods, respectively. Then the post-shock income can be written as $income_{it}^{post} = \sum_j (1 + z_s)e_{ist}$. One can obtain z_s from external information about sector-specific income shocks caused by the pandemic. For example, as previously stated, this paper uses the World Bank's report on the economic impacts of the pandemic in Vietnam.

[16] The rapid poverty reduction can be explained by a booming export sector and rising domestic demand (The World Bank, 2018).

[17] The rise in self-reported food insecurity might be due to changes in how the government classified poor households in 2016 (Decision 59/2015/QD-TTg).

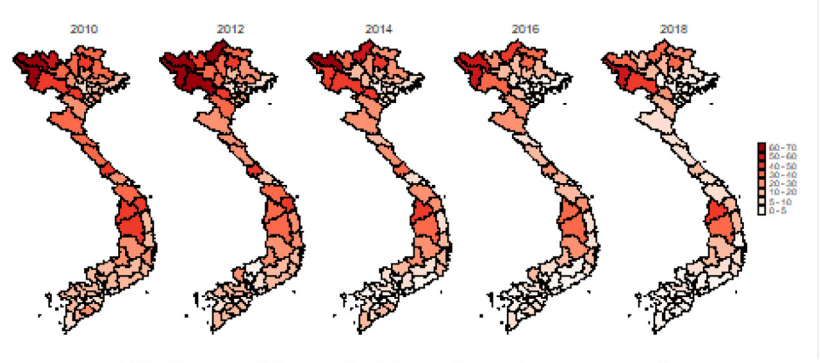
(a) Share of food-insecure household using self-reported data



(b) Share of food-insecure household using the HDDS



(c) Share of food-insecure household using staple calorie share



(d) Share of households below the poverty line

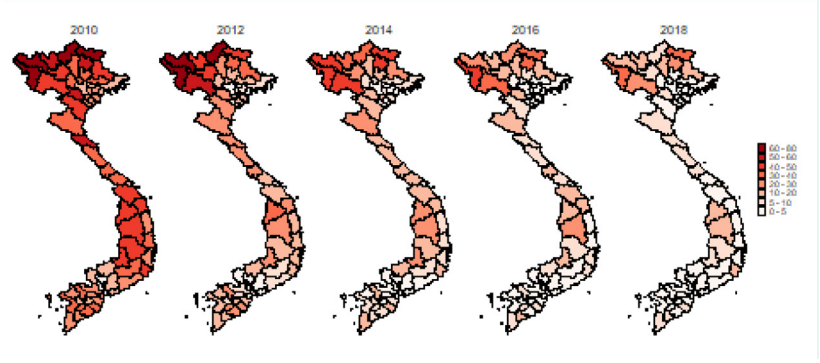


Fig. 3. Food insecurity and poverty by province and year.

The process to predict changes in food insecurity at some locality level involves two steps. In the first step, we use a regression to estimate β and θ on a household-level dataset and denote these estimates as $\hat{\beta}$ and $\hat{\theta}$; we refer to this as the "estimation" step. In the second step, we generate locality-level pre- and post-pandemic food insecurity risk and calculate the difference in order to measure changes in food insecurity; we refer to this as the "prediction" step. Specifically, we combine $\hat{\beta}$ and $\hat{\theta}$ with $income_{it}^{pre}, income_{it}^{post}$, and $\mathbf{Z_{it}}$ to generate the pre- and post-pandemic predicted probability of food insecurity at the household level:

$$\widehat{Pr}(Y_{it}=1)^{pre/post} = \hat{\beta}.income_{it}^{pre/post} + \mathbf{Z_{it}}\hat{\theta}, \tag{2}$$

where $\widehat{Pr}(Y_{it}=1)^{pre/post}$ is the predicted probability of food insecurity and reflects the "risk" that household i is food-insecure; that is, households with a higher value are more likely to be food insecure.

Although the predicted probability of food insecurity is a good indicator to measure food insecurity shock, it is perhaps more policy-relevant to predict the changes in the share of food-insecure households as a result of the income shock during the pandemic. Since our prediction step in Eq. 2 only provides a continuous variable for predicted probability of being food-insecure, we can only identify households with high risk of being food-insecure. By choosing a threshold c such that when the predicted probability exceeds the threshold, a household is classified as "high-risk". This additional step will then allow us to measure how the share of "high-risk" households changes due to the income shock.

Choosing a threshold to map the predicted probability values to a binary category, i.e. high-risk or not, is a standard step in the predictive modeling literature (James, Witten, Hastie, & Tibshirani, 2013). It is important to note that this threshold is applied to the *predicted probability of food insecurity* to classify a "high-risk" household. This classifying threshold is not the same as the cutoff applied to the *staple calorie share* to define a food-insecure household. The criterion to pick an optimal threshold for the predicted probability depends on the goal of the prediction. Our goal is to obtain the shares of "high-risk" households that is as close to the shares of households that are actually food-insecure; the first value is often known as the predicted prevalence while the latter is often known as the observed prevalence (Freeman & Moisen, 2008). Freeman and Moisen (2008) find that the two best approaches to satisfy this criterion. The first approach is choosing a threshold to minimize the difference between the predicted prevalence and the observed prevalence. The second approach is choosing a threshold to maximize the Cohen's Kappa statistic, which measures how close the predicted status and the actual status of food insecurity are, after accounting for the fact that they might be the same due to chance.[18]

Formally, let h_{it} denote the "high-risk" status of a household, i.e. the predicted food insecurity status, and is defined as $h_{it}^{c} = \mathbf{1}\left\{\widehat{Pr}(Y_{it}=1) > c\right\}$ while Y_{it} is the true food insecurity status. The predicted prevalence of threshold c, PP_c, is the probability of cases being predicted by threshold c as food-insecure: $PP_c = Pr(h^c = 1)$, and the observed prevalence (OP) is probability of a household being food-insecure, $OP = Pr(Y=1)$. The first approach is to select a threshold to minimize $d_c = |PP_c - OP|$ which is the difference between the predicted and observed prevalence rates.

The second approach is to choose a threshold with the largest Cohen's Kappa statistic, which is measured by

$$\kappa_c = \frac{Pr(h^c = Y) - Pr(h^c = Y | h^c \perp Y)}{1 - Pr(h^c = Y | h^c \perp Y)}$$

where $Pr(h^c = Y)$ is the probability that we correctly predict a case to be food-insecure or not, which measures the accuracy of the prediction. $Pr(h^c = Y | h^c \perp Y)$ is the probability that a case is correctly predicted when h^c and Y are independent of each other, which measures the accuracy by chance. Intuitively, the Cohen's Kappa measures the accuracy of the prediction after removing part of the accuracy that happens due to chance (Ben-David, 2007; Ben-David, 2008).

In the predictive modeling literature, the Cohen's Kappa can also be written in terms of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Specifically, the accuracy of the prediction is measured by the proportion of cases that are either true positive or true negative, $Pr(h^c = Y) = \frac{TP_c + TN_c}{N}$ where N is the sample size. The accuracy by chance is $Pr(h^c = Y | h^c \perp Y) = Pr(h^c = 1).Pr(Y = 1) + Pr(h^c = 0).Pr(Y = 0)$, which can be written as:

$$Pr(h^c = Y | h^c \perp Y) = \frac{(TP_c + FP_c)}{N} \cdot \frac{(TP_c + FN_c)}{N} + \frac{(TN_c + FN_c)}{N} \cdot \frac{(TN_c + FP_c)}{N}$$

We now turn to discuss the details of the estimation step. The simplest approach would be to use OLS to estimate the following linear probability model:

$$Y_{it} = \beta.income_{it} + \mathbf{Z_{it}}\theta + province_p \times year_t + \epsilon_{it}, \tag{3}$$

where Y_{it} is the binary variable for household's food insecurity status, Z_{it} denotes the time-varying household characteristics and $province_p \times year_t$ denote province-year fixed effects.[19] Two potential biases may arise with the OLS estimation. The first bias comes from unobserved factors that affect both income and food insecurity such as changes in household demographics or employment. The second bias is due to food insecurity simultaneously affecting household income by lowering productivity. Although controlling for household covariates and different levels of fixed effects may account for unobserved confounders, it is not sufficient to address the simultaneity problem.

We therefore use an instrumental variables (IV) approach to address the simultaneous bias. Specifically, household income can be instrumented by local exposure to national employment growth across different industries. As exposure to employment shocks affects household income (through household employment), households would change their food consumption accordingly. This is a form of a shift-share (or Bartik) instrument (Goldsmith-Pinkham, Sorkin, & Swift, 2020). The IV is constructed as follows:

$$IV_{dt} = \sum_j w_{bdj} \frac{(L_{tj} - L_{bj})}{L_{bj}},$$

where w_{bdj} denotes the employment share of industry j[20] in district d in the baseline period b, L_{tj} denotes the national employment of industry j in period t, and L_{bj} denotes the national employment of

[18] Another common approach in the machine learning prediction is choosing a threshold to maximize the true positive rate minus the false positive rate (also known as the Youden's J statistic). However, this approach is more appropriate for evaluating diagnostic test because it tends to overestimate the true prevalence when the event is rare (Freeman & Moisen, 2008).

[19] Controlling for province-year fixed effects is preferred for the purpose of estimation, but doing so will not generate parameter estimates for future years such as 2020. Therefore, controlling for province-linear trends is more suitable for the purpose of prediction (Auffhammer & Steinhauser, 2012; Newell, Prest, & and Sexton, 2018). In the next section, we show that both approaches yield similar estimates.

[20] For Vietnam, industries are classified using the three-digit ISIC system.

industry j in the baseline period b. Note that $\frac{(L_{tj}-L_{bj})}{L_{bj}}$ is the national employment growth rate of industry j in period t relative to the baseline period b. Our study period is between 2010 and 2018, and the baseline year is 2009. We estimate the following equation using a two-stage least squares (2SLS) approach:

$$\text{foodinsecurity}_{it} = \beta.\widehat{\text{income}}_{it} + \mathbf{Z_{it}}\theta + \mathbf{J_{d,2009}}\eta + province_p \times year_t + \epsilon_{it}, \quad (4)$$

and the first-stage equation is

$$\text{income}_{it} = \beta.\text{IV}_{dt} + \mathbf{Z_{it}}\zeta + \mathbf{J_{d,2009}}\alpha + province_p \times year_t + \epsilon_{it}, \quad (5)$$

where $\mathbf{J_{d,2009}}$ denotes district d's characteristics, which include demographics (gender and marriage), economics (wealth and unemployment), and education (college educated) in 2009 as control variables. This allows us to account for baseline factors that may affect both local exposure to and changes in food insecurity.

In a two-industry, one-period scenario, the instrument measures the variation of local exposure, measured by district-industry employment share in the baseline, to the national employment growth of one industry relative to another (Goldsmith-Pinkham et al., 2020). The research design thus captures the effect of the district-industry employment share on *changes* in food insecurity relative to the baseline period. In a multiple-period scenario, the effects of the baseline share on outcome are scaled by the national growth of employment. The research design is analogous to a dose–response design as we compare differential outcome changes in districts with different baseline shares of employment. Goldsmith-Pinkham et al. (2020) further show that in a multiple-industry, multiple-period scenario, the Bartik instrument is numerically equivalent to using multiple instruments that are baseline shares of different industries. The identifying assumptions for the estimates using a Bartik instrument to be consistent are that (1) the baseline district-industry share of employment is conditionally exogenous to changes in food insecurity, and (2) the baseline share only affects changes in food insecurity through the endogenous variable, which is household income (conditional on control variables).

We assess these assumptions in different ways. The IV exogeneity assumption is violated when there are other district-level characteristics in 2009 affecting both district-industry employment shares and changes in food insecurity. Therefore, we include control variables for various district-level demographic, economic, and educational characteristics in 2009 in Eq. 4 and 5. These baseline characteristics can explain up to 94% of the variation in the baseline industry-district employment share (see Table 4), which means controlling for these factors will account for most of the endogeneity concerns with the IV (Goldsmith-Pinkham et al., 2020). Furthermore, we also follow the recommendations in Goldsmith-Pinkham et al. (2020) by estimating the IV model using baseline industry-district employment shares as multiple IVs to employ an overidentification test. That is, we estimate Eq. 3 where the first-stage equation is

$$\text{income}_{it} = \sum_t \sum_j \gamma_{jt}.(w_{bdj} \times year_t) + \mathbf{Z_{it}}\zeta + \mathbf{J_{d,2009}}\alpha + province_p \times year_t + \epsilon_{it} \quad (6)$$

and $w_{bdj} \times year_t$ is the industry-district employment share in 2009 times the year dummy variables. Because there are multiple instruments, we use different estimators including generalized method of moments (GMM) and limited information maximum likelihood (LIML) since they are more appropriate when dealing with many instruments. We use the Sargan–Hansen overidentification test for the 2SLS estimator and the Anderson-Rubin overidentification test for the GMM and LIML estimators.

The exclusion restriction assumption is not testable in this study, but it likely holds because employment shocks can only affect household food insecurity through changes in household income. There are two potential concerns about the exclusion restriction assumption: (1) the employment shocks may shift the market demand for food and change the market food price, which in turn affects household food consumption decisions; and (2) the employment shock in the agriculture sector may affect agricultural wages, which in turn will also affect food prices. We account for local market conditions including food prices by controlling for province-year fixed effects. We also check for potential violations; we construct a "non-agriculture Bartik IV" where we only consider non-agriculture sectors for the shift-share IV, and, in the next section, we compare the results using this IV with the main results.

Our approach has the following caveats. First, the estimation step uses a linear probability model, so the predicted probability is not bounded between 0 and 1; this feature is undesirable compared to standard binary dependent variable models such as the logistic or probit regression. An alternative approach is using an IV probit model which will provide predicted probability of food insecurity that is bounded between 0 and 1. However, this control function estimator is not preferred because it requires that the first stage model is correctly specified. In other word, if we do not include the correct set of instruments, the control function estimator is no longer consistent (Lewbel, Dong, & Yang, 2012).[21] In contrast, a true instrumental variable estimator (such as the linear IV approach that we use) only requires that the instrumental variable is correlated with the endogenous regressor and uncorrelated with the error term; it does not impose any structural assumptions on the errors in the first stage regression. When we do not include all of the right instruments in the first stage, the estimator is less efficient but still consistent (Lewbel et al., 2012). Therefore, making prediction based on an IV probit approach will lead to the same problem as the OLS approach unless we know the correct set of instruments in the first stage. In Section 5.2, we show that our linear IV approach actually outperforms the IV probit approach in out-of-sample prediction despite the unbounded predicted probability problem.

More importantly, since our prediction approach relies on income shocks induced by changes in the labor market for the IV estimation, it implicitly assumes that such shocks are similar to the income shocks caused by the pandemic. If the two shocks are very different from one another, then the IV-based prediction may no longer be accurate. Specifically, the income shocks from our IV estimates are broad changes in income due to national labor demand shocks, while the income shocks caused by the pandemic might be driven by changes in employment, especially more vulnerable workers such as those in the informal sector, part-time employees, or workers with lower education or less experience. We compare the prediction from our approach with the food insecurity measure generated from the actual income data during the pandemic in Appendix D and find that our prediction is highly accurate, suggesting that a violation of this assumption is not a major empirical concern.

Another caveat of our approach is that the household income function does not include non-working income such as remittances and social transfers, which can play a major role in shielding households from food insecurity. Although it can easily be extended to account for such features, we choose not to include these variables because such data are not always available, especially in labor force surveys. External information about changes in remittance is also unlikely to be available. Therefore, we interpret our prediction as a lower bound of the actual food insecurity

[21] This is true even when the IV probit is estimated using maximum likelihood.

shock when remittances fall as they did during this pandemic (The World Bank, 2020).

4. Estimating the effect of household income on food insecurity

In this section, we present the results from estimating the relationship between household income and food insecurity using the 2010–2018 VHLSS data for the OLS and IV approaches. Specifically, we use OLS to estimate Eq. 3 and 2SLS to estimate Eq. 4 and 5 and report the results in Section 4.1. We then discuss different validity assessments in Section 4.2.

4.1. Regression results

Table 2 provides the regression estimates for the effect of household income on food insecurity. Panel A reports the estimates from the OLS approach, and panel B reports the estimates from the IV approach (the estimates for the first-stage equation are reported in Table A.1). We consider controlling for different levels of fixed effects to account for any unobserved heterogeneity. For the models in columns 1, 2, and 3, we alternatively control for province, district, and household fixed effects. The household fixed effects model is run on a subsample that forms an unbalanced household panel, which is why there are fewer observations. The model in column 4 controls for province and year fixed effects, and the model in column 5 also controls for province-specific linear trends. In column 6, the model controls for province-by-year fixed effects.

All models control for household characteristics: urban, household size, and the fraction of households with postsecondary education. We also control for district demographic and economic characteristics in 2009 that might correlate with food insecurity and district-industry employment shares in 2009. These characteristics include gender, marital status, college degree holders, disability, immigration status, average wealth,[22] and unemployment rate. In the household fixed effects model, these time-invariant control variables are not included. For the OLS estimations, we cluster the standard errors at the commune-year level to account for the survey's sampling design. In the IV estimation, we cluster the standard errors at the district level to account for district-level exposure to the labor market shocks (Abadie, Athey, Imbens, & and Wooldridge, 2017).

We first compare the estimates from the province fixed effects, district fixed effects, and household fixed effects models in columns 1 to 3. The IV estimates are higher than the OLS estimates for all specifications. In column 1, the OLS estimate is –0.124 and the IV estimate is –0.263 when controlling for province fixed effects. In column 2, the estimates are –0.111 for OLS and –0.222 for IV when controlling for district fixed effects. In column 3, the estimates are –0.071 for OLS and –0.188 for IV when controlling for household fixed effects (note that this is estimated on a subsample that forms an unbalanced household panel).

Given that the estimates do not vary considerably across different levels of fixed effects, we focus on models that control for province fixed effects and secular trends. Because the variation of the shift-share IV is at the district-year level, controlling for district fixed effects or household fixed effects and trends would absorb all of this variation. We compare the models that control for province fixed effects and year fixed effects (column 4), province fixed effects and province-specific linear trends (column 5), and province, year, and province-by-year fixed effects (column 6). In column 4, the estimates are –0.113 for the OLS approach and 0.182 but are insignificant for the IV approach. In column 5, the estimates are –0.112 for OLS and –0.349 for IV. In column 6, the estimates are –0.112 for OLS and –0.396 for IV. The specifications in column 4 only control for trends at the national level, while the specifications in columns 5 and 6 control for trends at the province level. It appears that controlling for province-linear trends and controlling for province-by-year fixed effects yield similar estimates. These results strongly suggest that the main specification used in the prediction is correctly specified.

Our findings are generally consistent across different measures of food insecurity. We estimate our models using the self-reported and HDDS-based food insecurity outcomes. The results are reported in Table 3. As mentioned before, using self-reported outcomes severely underestimates the effect of income on food insecurity because only poor households are asked to report their food insecurity status; the estimates also have reverse signs when controlling for time trends or year fixed effects. The results for the HDDS-based measure are qualitatively similar to our main findings.

4.2. Assessing the validity of the shift-share instrument

To assess the validity of the IV approach, we first assess the exogeneity assumption of the shift-share IV. Following Goldsmith-Pinkham et al. (2020), we identify district-level factors that are correlated with the initial industry shares of employment in 2009. We consider the district-level wealth index; the population shares of those who are female, married, college educated, disabled, and an immigrant; and the unemployment rate, which is calculated on the sample of 15-to-69-year-olds using the 2009 census data. Table 4 presents the estimates from regressions of these covariates on the initial industry share of labor in 2009. It is implied from the R-squared that the covariates explain mostly 35%-94% of the 2009 industry share of employment variation. We control for these variables to avoid model misspecification due to omitted variables.[23]

Given that the 2SLS estimator with the Bartik IV is numerically equivalent to the GMM estimator that uses industry shares of employment as instruments, we test for the validity of Bartik's instrument by estimating a regression with multiple IVs as initial industry-district shares of employment times year dummy variables and conduct an overidentification test on this regression, as suggested in Goldsmith-Pinkham et al. (2020). That is, we estimate Eq. 3 where the household income is instrumented as indicated in Eq. 6. Table 5 presents the results. We note that (1) our overidentified IV estimations are very similar to the main findings, and (2) our results fail to reject the overidentification tests after including the district-level factors as controls. This evidence suggests that the main model is correctly specified.

Second, we assess the exclusion restriction assumption by estimating using non-agriculture Bartik IV. Table 6 presents the results. The estimates using the alternative Bartik IV are very similar to the estimates from the main results. This suggests that the agriculture component of the Bartik IV does not independently affect our main results.

5. Predicting food insecurity risks in Vietnam

In this section, we apply the prediction approach proposed in Section 3.3 to make predictions about changes in food insecurity risk in Vietnam. First, we use the 2010–2018 VHLSS to select the optimal classifier threshold in Section 5.1. As discussed in Section 3.3, we need to select a threshold to classify the high-risk" households based on the predicted probability of food insecurity. In Section 5.2, we assess the out-of-sample predictive accuracy of

[22] Wealth is estimated using principal component analysis on electricity, piped water, air conditioner, computer, washing machine, refrigerator, television, and radio.

[23] See Goldsmith-Pinkham et al. (2020) for a detailed discussion on this approach.

Table 2
Estimates for income effects on food insecurity.

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS approach						
	−0.124***	−0.111***	−0.071***	−0.113***	−0.112***	−0.112***
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.003)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
N	46443	46443	27657	46443	46443	46443
Panel B: Bartik IV approach						
	−0.263***	−0.222***	−0.202***	0.182	−0.349***	−0.396***
	(0.018)	(0.015)	(0.026)	(0.231)	(0.132)	(0.130)
	[0.000]	[0.000]	[0.000]	[0.430]	[0.008]	[0.002]
Cragg-Donald F-stat	381.86	372.58	632.82	4.54	6.99	7.94
Montiel Olea and Pflueger F-stat	2.0e + 05	1.9e + 05		2336.57	3599.18	4087.31
N	46443	46443	24499	46443	46443	46443
Additional controls						
Province FE	✓			✓	✓	✓
District FE		✓				
Household FE			✓			
Year FE				✓		✓
Province FE x linear trends					✓	
Province x Year FE						✓

The table report results from OLS and IV estimations for the effect of household income on food insecurity in Eq. 3. IV estimation instruments household income with Bartik IV (see Eq. 5). All models control for urban, household size, and the fraction of households with postsecondary education, and 2009 district characteristics (gender, marital status, college degree, immigration status, disability, average wealth index, and unemployment rate). The household fixed effects model does not include 2009 controls. Standard errors are reported in parentheses, and p-values are reported in brackets. OLS standard errors are clustered at the commune-year, IV standard errors are clustered at the district level, and household FE estimations are estimated on a subsample of households that form an unbalanced panel with robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

our IV-based approach and the standard OLS-based approach by comparing the predictions for 2016 and 2018 with the actual food insecurity data for these two years. In Section 5.3, we predict food insecurity changes due to income shocks during the pandemic for 702 districts in Vietnam. In Section 5.4, we discuss how these results can inform policymakers and international organizations about where to prioritize assistance.

5.1. Choosing the optimal classifier threshold

Following the estimation step, we generate the predicted probability of food insecurity on the 2010–2018 VHLSS data and plot the distributions by their actual food insecurity status in Fig. A.1. We choose the optimal threshold c^* between 0th and 99th percentile given that our regression is a linear probability model.

In Fig. A.2, we plot the difference between the observed and predicted prevalence rates as well as the Cohen's Kappa statistics for $\forall c \in [0, 99]$. As discussed in Section 3.3, the threshold with the smallest difference in prevalence line satisfies the first criterion, while the threshold at the highest point of the Cohen's Kappa curve satisfies the second criterion. We find that the 85th percentile satisfies both criteria. Therefore, we classify all households with predicted probability of food insecurity above the 85th percentile as "high-risk" households for the rest of the paper.

5.2. Out-of-sample prediction validation in Vietnam

To validate the process described in Section 3.3, we will predict provinces' share of households with high food insecurity risk in 2016 using the 2016 LFS data and compare with provinces' share of households that are actually food-insecure in 2016 from the 2016 VHLSS data.[24] Specifically, we implement the estimation step using the 2010–2014 VHLSS data, and the prediction step on the 2016 LFS data. We also conduct another validation exercise by predicting the 2018 food insecurity using the LFS data and compare with the actual food insecurity data from the 2018 VHLSS data. The OLS-based prediction is used as a benchmark because it is commonly used to predict food insecurity (Lentz et al., 2019; Gundersen et al., 2020; Schanzenbach & Pitts, 2020).

To quantify the difference between our predictions with the actual data, we calculate the sum of squared errors between the two measures; that is, $\sum_p e_p^2 = \sum_p \left(f_p^{actual} - f_p^{predicted}\right)^2$, where f_p measures the province-level share of households with actual food insecurity or the high predicted risk for province p, as a formal measure of predictive accuracy (Auffhammer & Steinhauser, 2012; Athey, 2018). Another measure of predictive performance is the R-squared when regressing the predicted share against the actual share of food-insecure households, which reflects how much variance of the actual share is correctly predicted by each approach. We also plot these province-level percentages along with a 45-degree line in Fig. 4. The x-axis shows the share of households with actual food insecurity, and the y-axis shows the percentage of households with high predicted risk. The higher the dot compared to the 45-degree line, the more we overestimate food insecurity in that province; the lower the dot, the more we underestimate it. In other words, a model with dots closer to the red line has higher predictive accuracy. We report the results for 2016 in (a) and 2018 in (b).

We find that IV-based prediction outperform the OLS-based predictions in both 2016 and 2018. Graphically, most points from the IV models are relatively close to the 45-degree line, while most points in the OLS models are much higher than the line. The sums of squared errors are 2.55 for OLS-based and 0.69 for IV-based predictions in 2016. In 2018, the sums of squared errors are 2.61 for OLS-based and 0.79 for IV-based prediction. Therefore, the IV models are far more accurate than the OLS models. Similarly, IV models have a higher R-squared in both years relative to OLS models, indicating that the IV-based prediction can explain more of the actual

[24] For these validation exercises, we use province-level shares instead of district-level shares because the VHLSS data are not representative at the district level. We aggregate the predicted data (from the LFS data) at the province level to be comparable with the VHLSS data.

Table 3
Estimates for income effects on alternative food insecurity measures

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Self-reported food insecurity outcome						
OLS approach						
	−0.039***	−0.034***	−0.014***	−0.037***	−0.037***	−0.037***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
N	46443	46443	27709	46443	46443	46443
IV approach						
	−0.051***	−0.057***	−0.050***	0.204	0.081	0.053
	(0.008)	(0.006)	(0.014)	(0.144)	(0.083)	(0.064)
	[0.000]	[0.000]	[0.000]	[0.157]	[0.330]	[0.406]
Cragg-Donald F-stat	381.86	372.58	631.13	4.54	6.99	7.94
Montiel Olea and Pflueger F-stat	2.0e + 05	1.9e + 05		2336.57	3599.18	4087.31
N	46443	46443	24570	46443	46443	46443
Panel B: HDDS-based food insecurity outcome						
OLS approach						
	−0.126***	−0.106***	−0.067***	−0.124***	−0.124***	−0.124***
	(0.003)	(0.003)	(0.008)	(0.004)	(0.004)	(0.004)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
N	46443	46443	27707	46443	46443	46443
IV approach						
	−0.242***	−0.134***	−0.118***	−1.271**	−1.027***	−1.012***
	(0.026)	(0.019)	(0.034)	(0.525)	(0.330)	(0.304)
	[0.000]	[0.000]	[0.000]	[0.016]	[0.002]	[0.001]
Cragg-Donald F-stat	381.86	372.58	669.48	4.54	6.99	7.94
Montiel Olea and Pflueger F-stat	2.0e + 05	1.9e + 05		2336.57	3599.18	4087.31
N	46443	46443	24587	46443	46443	46443
Additional controls						
Province FE	✓			✓	✓	✓
District FE		✓				
Household FE			✓			
Year FE				✓		✓
Province FE x linear trends					✓	
Province x Year FE						✓

The table reports results from OLS and IV estimations for the effect of household income on food insecurity in Eq. 3. IV estimation instruments household income with Bartik IV (see Eq. 5). All models control for urban, the household size, the fraction of households with a postsecondary education, and 2009 district characteristics (gender, marital status, college degree, immigration status, disability, average wealth index, and unemployment rate). The household fixed effects model does not include 2009 controls. Standard errors are reported in parentheses, and p-values are reported in brackets. OLS standard errors are clustered at the commune-year, IV standard errors are clustered at the district level, and household fixed effects estimations are estimated on a subsample of households that form an unbalanced panel with robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

variation in share of food insecurity. These results suggest that the IV approach has higher out-of-sample predictive accuracy in predicting food insecurity.

We also apply these validation steps for the IV probit approach and compare with the results for the linear IV approach. In other word, we estimate the relationship between household income and food insecurity status using IV probit and generate provinces' share of household with high food insecurity risk for 2016 and 2018.[25] We then validate these predictions with the actual shares of food-insecure households for these years using the VHLSS data. We find that the IV probit approach yields more accurate prediction than the OLS approach, but less accurate than our linear IV approach for both 2016 and 2018 (see Fig. A.4). As explained in Section 3.3, the assumption of the IV probit approach is relatively strong and unlikely to be met in this study. Therefore, it is not surprising that the linear IV approach outperforms the IV probit approach.

5.3. Predicting changes in food insecurity of Vietnam during the pandemic

We use the 2019 Labor Force Survey (LFS) data to predict the share of households with high risk of food insecurity at the district level before and after the income shock using the method described in Section 3.3. We first obtain the pre-pandemic predicted probability of food insecurity for each household in the 2019 LFS sample using Eq. 2, where the coefficients come from the IV estimates in Table 2 and the right-hand side variables including household income are from the LFS data.[26] We then classify those with the predicted probability above the 85th percentile as "high-risk" households and obtain pre-pandemic share of "high-risk" households at the district level.

Post-pandemic predicted probability of food insecurity can be generated in the same way except for using the post-pandemic

[25] We use the same criterion to find the optimal threshold for the predicted probability of food insecurity generated from the IV probit estimation. The threshold that minimizes the difference between the predicted prevalence and observed prevalence is 33%, and the threshold that maximizes the Cohen's Kappa statistic is 35% (see Fig. A.3). We choose the 33% as the threshold for the predicted probability; the results are similar for the 35% threshold.

[26] Recall that the LFS data itself do not contain information about food insecurity.

Table 4
Relationship between industry-district employment shares and district-level characteristics in 2009.

Outcome	Agriculture (1)	Mining (2)	Manufact. (3)	Utility (4)	Construction (5)	Retail (6)	Logistics (7)	Hospitality (8)	Media (9)	Finance (10)
Urban	−0.20***	0.00***	0.04**	0.00***	0.01**	0.04**	0.02**	0.03**	−0.00***	0.00***
	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Wealth	−0.21***	0.02**	0.07*	0.00***	0.01**	0.06*	0.02**	0.03**	−0.00***	0.00***
	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Female share	−2.22***	−0.33***	1.38	−0.02***	0.53	0.37	0.01**	0.14	0.01***	0.04**
	(0.40)	(0.12)	(0.24)	(0.02)	(0.12)	(0.12)	(0.05)	(0.06)	(0.01)	(0.01)
Married share	−0.29***	0.11	0.15	0.03**	0.19	−0.10***	−0.02***	−0.11***	0.01***	0.00***
	(0.16)	(0.05)	(0.09)	(0.01)	(0.05)	(0.05)	(0.02)	(0.02)	(0.00)	(0.00)
College+	0.62	−0.12***	−0.91***	0.01**	−0.10***	−0.16***	−0.07***	−0.11***	0.12	0.10
	(0.13)	(0.04)	(0.08)	(0.01)	(0.04)	(0.04)	(0.02)	(0.02)	(0.00)	(0.00)
Disability	−0.43***	0.05*	−0.49***	−0.05***	0.74	0.01**	0.03**	0.05**	0.04**	0.02**
	(0.71)	(0.22)	(0.42)	(0.03)	(0.21)	(0.21)	(0.09)	(0.10)	(0.02)	(0.02)
Immigration status	−1.22***	0.05*	1.03	0.02**	0.14	0.08*	0.02**	0.01**	−0.02***	−0.01***
	(0.10)	(0.03)	(0.06)	(0.00)	(0.03)	(0.03)	(0.01)	(0.01)	(0.00)	(0.00)
Uemployed	−4.45***	0.26	0.59	0.09*	0.67	1.27	0.60	0.50	0.00***	−0.00***
	(0.51)	(0.16)	(0.30)	(0.02)	(0.15)	(0.15)	(0.07)	(0.08)	(0.01)	(0.01)
R-squared	0.92	0.35	0.79	0.69	0.58	0.88	0.85	0.87	0.92	0.94
Observations	703	703	703	703	703	703	703	703	703	703

Outcome	Real Estate (11)	Science (12)	Admin (13)	Govern. (14)	Education (15)	Health (16)	Entertain. (17)	Other Services (18)	Household workers (19)
Urban	−0.00***	−0.00***	−0.00***	0.02**	0.02**	0.01***	0.01***	0.01***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Wealth	0.00***	−0.00***	0.00***	−0.01***	−0.01***	0.00***	0.00***	0.01***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female share	0.01***	0.02**	0.02**	−0.11***	−0.02***	0.03**	0.07*	0.03**	0.04**
	(0.01)	(0.01)	(0.01)	(0.05)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Married share	−0.01***	−0.01***	−0.01***	0.11	−0.00***	0.01***	−0.01***	−0.03***	−0.02***
	(0.00)	(0.01)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
College+	0.01**	0.12	0.02**	0.23	0.22	0.05**	−0.01***	−0.03***	0.02**
	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Disability	0.01***	0.03**	0.01**	−0.10***	−0.09***	−0.03***	0.04**	0.10*	0.06*
	(0.01)	(0.02)	(0.01)	(0.09)	(0.07)	(0.03)	(0.03)	(0.03)	(0.02)
Immigration status	0.02**	−0.01***	0.01***	−0.04***	−0.06***	−0.02***	0.00***	0.01***	−0.01***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Uemployed	0.02**	0.02**	0.02**	0.12	0.09*	0.05*	0.01**	0.15	0.01**
	(0.01)	(0.02)	(0.01)	(0.07)	(0.05)	(0.02)	(0.02)	(0.02)	(0.02)
R-squared	0.76	0.90	0.88	0.67	0.75	0.77	0.66	0.89	0.72
Observations	703	703	703	703	703	703	703	703	703

The table reports results from estimating a regression of industry-district shares of employment (as outcome) and district-level characteristics in 2009. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5
IV estimation on the effects of household income on food insecurity using 2009 district-industry employment shares × year dummy as instruments.

	2SLS		GMM		LIML	
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	−0.254***	−0.365***	−0.290***	−0.386***	−0.283***	−0.397***
	[-0.31; −0.20]	[-0.44; −0.29]	[-0.34; −0.24]	[-0.46; −0.31]	[-0.35; −0.21]	[-0.49; −0.31]
	(0.029)	(0.040)	(0.025)	(0.039)	(0.036)	(0.047)
Overidentification test						
J Statistics	53.17	22.64	53.17	22.64	48.98	21.33
χ-squared p-value	0.000	0.161	0.000	0.161	0.000	0.212
Obs	46443	46443	46443	46443	46443	46443
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Province-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline district control	No	Yes	No	Yes	No	Yes

The table presents the results from estimating the effects of household income on food insecurity in Eq. 3, where household income is instrumented with industry-district employment shares in 2009 × year dummy variables; the first-stage equation is in Eq. 6. Household characteristic controls include urban, household size, and share of household members with a postsecondary education. Baseline (2009) district-level characteristics are gender, marital status, college degree, immigration status, disability, average wealth index, and unemployment rate. The Sargan–Hansen test is the overidentification test for 2SLS, and the Anderson-Rubin overidentification test is for GMM and LIML. Standard errors are clustered at the district level. Standard errors are reported in parentheses, and p-values are reported in brackets. $***p < 0.01, **p < 0.05, *p < 0.1$.

Table 6
Estimates for income effects on food insecurity using non-agriculture Bartik IV.

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Bartik IV						
	−0.263***	−0.222***	−0.202***	0.182	−0.349***	−0.396***
	(0.018)	(0.015)	(0.026)	(0.231)	(0.132)	(0.130)
	[0.000]	[0.000]	[0.000]	[0.430]	[0.008]	[0.002]
N	46443	46443	24499	46443	46443	46443
Panel B: Non-ag. Bartik IV						
	−0.240***	−0.195***	−0.188***	0.567	−0.238**	−0.326**
	(0.017)	(0.013)	(0.018)	(0.519)	(0.100)	(0.130)
	[0.000]	[0.000]	[0.000]	[0.275]	[0.017]	[0.012]
N	46443	46443	24499	46443	46443	46443
Additional controls						
Province FE	✓			✓	✓	✓
District FE		✓				
Household FE			✓			
Year FE				✓		✓
Province FE x linear trends					✓	
Province x Year FE						✓

The table report results from IV estimations for the effect of household income on food insecurity in Eq. 3, where household income is instrumented with Bartik IV and non-agriculture Bartik IV (see Eq. 5). All models control for urban, household size, and the fraction of households with a postsecondary education, and 2009 district characteristics (gender, marital status, college degree, immigration status, disability, average wealth index, and unemployment rate). Standard errors are reported in parentheses, and p-values are reported in brackets. The household fixed effects model does not include 2009 controls. OLS standard errors are clustered at the commune-year, IV standard errors are clustered at the district level, and household fixed effects estimations are estimated on a subsample of households that form an unbalanced panel with robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

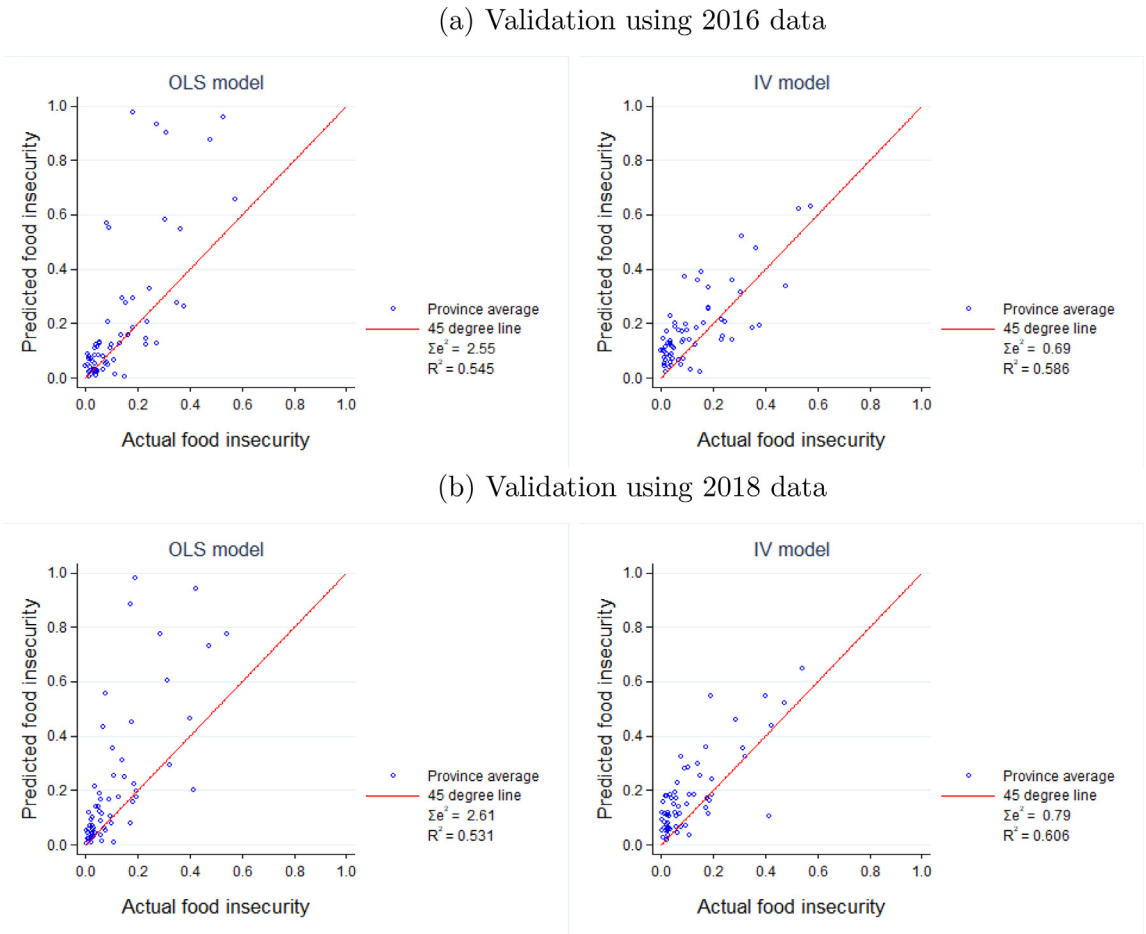


Fig. 4. Predicted and actual food insecurity of 2018 at the provincial level. Note: *Predicted food insecurity* measures each province's share of households with high predicted food insecurity risk (above 84th percentile); *Actual food insecurity* measures the share of households with actual food insecurity in each province using the VHLSS sample in year *y*. Predicted food insecurity risk is obtained by estimating the effect of income on food insecurity using the pre-*y* VHLSS sample and predicting using the LFS sample in *y*.

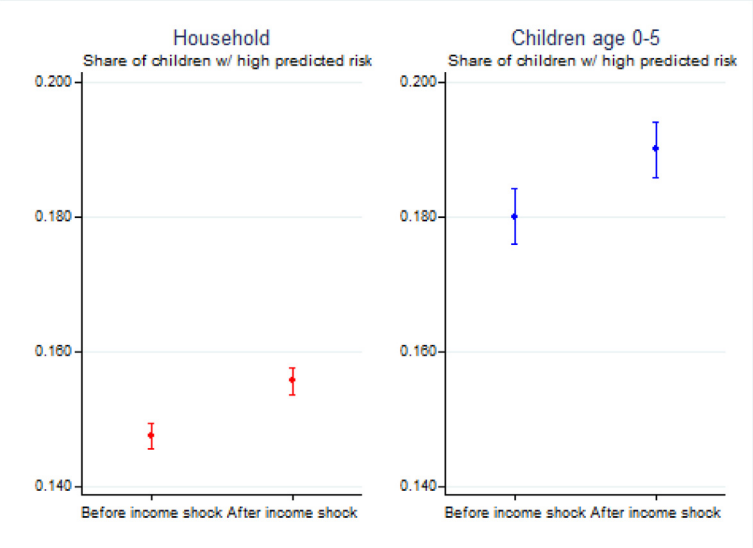


Fig. 5. Percentage of households with high predicted risk.

income instead of the 2019 income. The World Bank indicates that the average monthly income per person in the agricultural, manufacturing, and service sectors in quarter 2 of 2020 is lower than that of 2019 by 3%, 5.1%, and 7.3%, respectively (Morisset et al., 2020). [27] We use these estimates to calculate the post-pandemic income for each household in the 2019 LFS sample,[28] and generate the post-pandemic predicted probability for each household in the 2019 LFS sample following Eq. 2. Similarly, "high-risk" households are classified using the 85th percentile threshold to obtain the post-pandemic share of high-risk households for each district.

We find that the average share of "high-risk" household is 14.74% for before pandemic and 15.56% for after pandemic. In other word, the share of food-insecure households is predicted to increase by 0.82 percentage points (95% CI: 0.77, 0.87) due to the income shock during the pandemic. Next, we turn to the most vulnerable population, young children: the average share of children ages 0 to 5 with high food insecurity risk is 18% for before pandemic and 19% for after pandemic. The share of food-insecure children is predicted to increase by 0.997% (95% CI: 0.897, 1.10). These results are summarized in Fig. 5.

This small increase in food insecurity, however, masks significant geographic variation in changes in food insecurity across districts. In Fig. 6, we map the districts' share of households with high pre- and post-pandemic risk and the difference between the two shares. The increases were relatively small across the country, but a small number of districts experienced an increase as large as 7.86 percentage points. Only 102 out of 702 districts in our sample experienced an increase larger than 2 percentage points. Similarly, Fig. 7 shows the changes in the percentage of children ages 0 to 5 with a high predicted food insecurity risk due to the income shock; 560 districts are predicted to have a increase between 0 and 2 percentage points, 141 districts are predicted to have an increase between 2 and 10 percentage points, and only 7 districts are predicted to have an increase larger than 10 percentage points and up to 19.33 percentage points. This finding suggests that children might be affected more severely in terms of food insecurity during the pandemic.

Alternatively, we can ask what the food insecurity risks would have been if Vietnam had not contained the virus successfully. To answer this question, we predict food insecurity for a scenario where the quarterly reductions from Morisset et al. (2020) were to continue for all four quarters. In other words, we calculate the annualized changes of income and use those calculations to make our food insecurity prediction. Under this scenario, the average income decreases in the agricultural, manufacturing, and service sectors by 12.6%, 22% and 32.55%, respectively. The share of "high-risk" households before the pandemic is 12.46% and the share after the pandemic is 19.28%. In other word, the share of food-insecure household is predicted to rise by 6.81 percentage points (95% CI: 6.67, 6.95). In Fig. A.5, we map the changes of the shares of "high-risk" households due to the income shock during the pandemic. A similar pattern is expected: while the average change is 7.06 percentage points, a small number of districts would have experienced an increase of up to 26.62 percentage points. This is likely an underestimation because those who are food insecure in the short run may experience a decrease in productivity in the long run, which in turn affects both long-run income and food insecurity.

5.4. Policy implications

These predictions provide a detailed picture of which districts may have experienced larger increases in food insecurity and hence may require more assistance than others. This information will allow the government or international organizations to quickly identify and allocate more resources toward districts experiencing a larger increase in food insecurity risk and allocate fewer resources toward districts with a smaller increase in risk. As an example, consider Fig. A.6 scatter plot of each district's pre-pandemic share of households with high food insecurity risk and each district's percentage point increase in such share due to the pandemic. Districts on the right of the dashed line are in the top 10% in terms of food insecurity, i.e. districts with the largest shares of "high-risk" households in 2019, and districts above the solid line

[27] These figures are based on the GSO's estimation using the LFS data for the second quarter of 2020. According to the GSO report, the average monthly income reduces from 5,517 VND in Q2-2019 to 5,238 VND in Q2-2019 (5.06%). Income reduces from 3,035 VND to 2,951 VND (2.76%) in the agriculture, forestry, and fishery sector; from 6,534 VND to 6,201 VND (5.01%) in the industry and construction sector; and from 6,939 VND to 6,429 VND (7.35%) in the service sector.

[28] Note that the income and employment data from the LFS sample are classified at the industry level (three-digit ISIC), while the information about the income shocks is by sector. Therefore, we re-classify the income and employment data by sector before matching with the sector-specific income shocks.

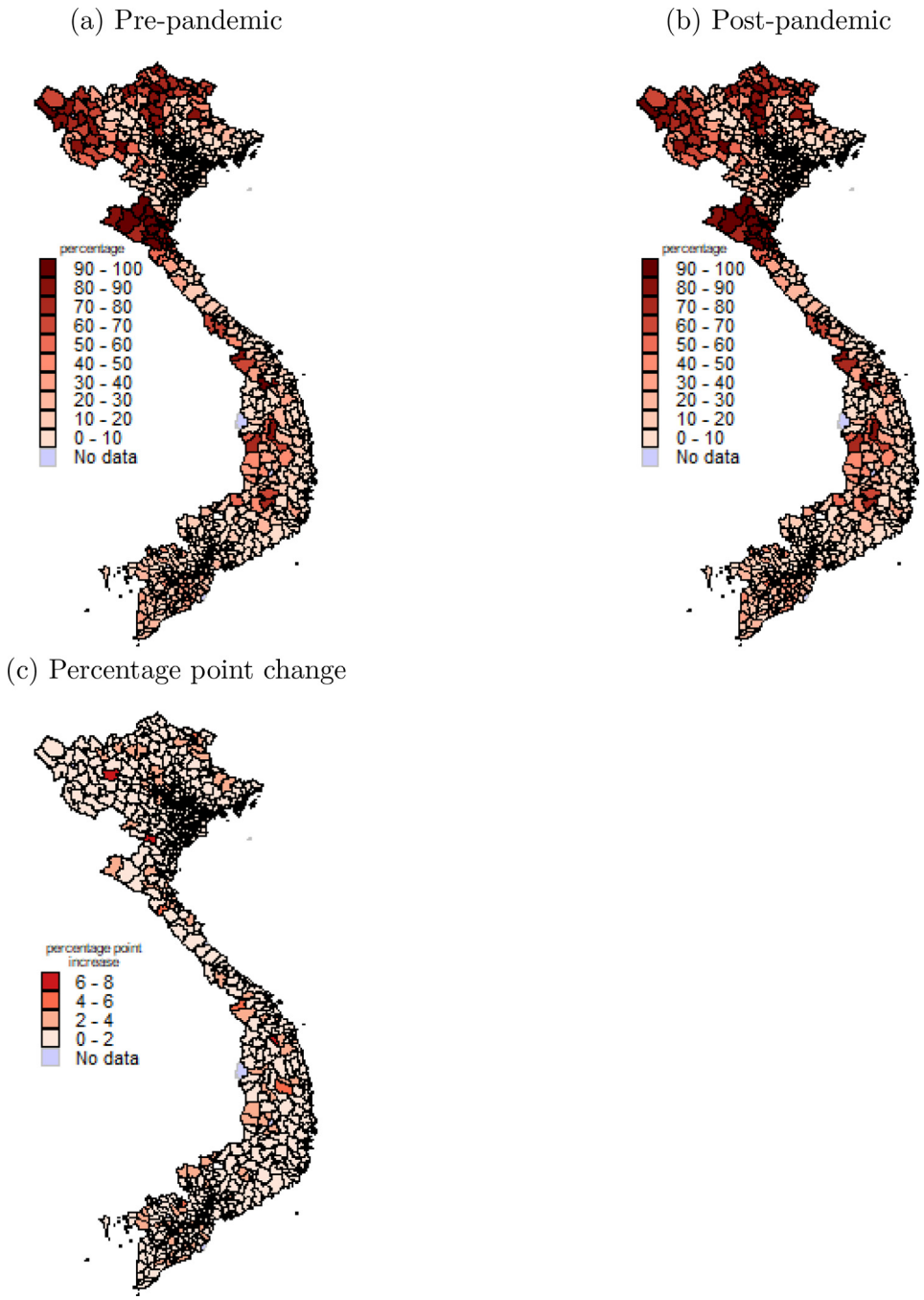


Fig. 6. Pre-pandemic, post-pandemic, and percentage point change in each district's share of households with high food insecurity risk. The maps show the (a) pre-pandemic, (b) post-pandemic, and (c) percentage point change in the share of households with high food insecurity risk. Pre-pandemic risk is predicted food insecurity using 2019 household income data from the 2019 Labor Force Survey. Post-pandemic risk is predicted food insecurity using post-pandemic income, which is calculated using the quarterly percentage change of average income by sector from Morisset et al. (2020): average income decreased by 3% in agriculture, 5.1% in manufacturing, and 7.3% in services per quarter.

in the top 10% in terms of predicted increase in food insecurity, i.e. districts with the largest increase in the share of "high-risk" households due to the pandemic.

One potential way that the government can allocate aid based on needs is prioritizing districts that are in the first quadrant since they both have substantially high pre-pandemic food insecurity and the largest increase in food insecurity due to the pandemic. The second priority can be assigned to districts to the left of the dashed line and above the solid line; these are districts that experience a temporary surge in food insecurity. The third priority can be of districts to the bottom right corner; although these districts have a substantially high pre-pandemic food insecurity, the increase is not as high compared to other districts. More importantly, their pre-pandemic level of food insecurity can be linked to factors unrelated to affordability, such as a traditional diet. The districts in the bottom left of the graph have the lowest risk and hence would be assigned the lowest priority. The government can also pursue a more sophisticated prioritization scheme based on these food insecurity risks, for example, employing multiple cutoffs.

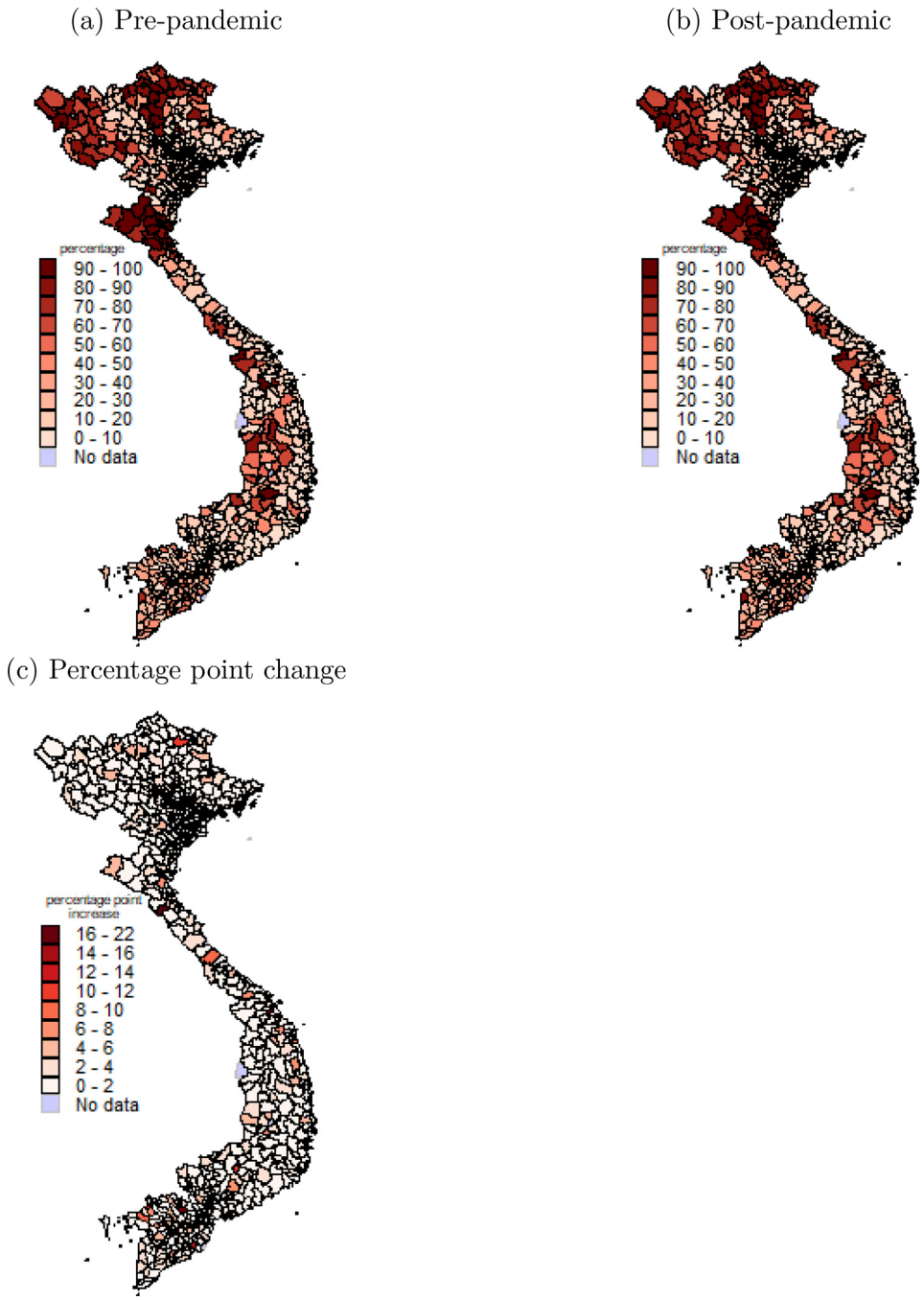


Fig. 7. Pre-pandemic, post-pandemic, and percentage point change in each district's share of children ages 0–5 with high food insecurity risk. Maps show the pre-pandemic, post-pandemic, and percentage point change in the share of children ages 0–5 with a high food insecurity risk. Pre-pandemic risk is predicted food insecurity using 2019 household income data from the 2019 Labor Force Survey. Post-pandemic risk is predicted food insecurity using post-pandemic income, which is calculated using the quarterly percentage change of average income by sector from Morisset et al. (2020): average income decreased by 3% in agriculture, 5.1% in manufacturing, and 7.3% in services per quarter.

It is important to note that an effective targeting approach would combine the geographical targeting approach proposed in this study with household or individual targeting (Barrett, 2010; World Food Programme, 2015). The method proposed in this study can first provide a district-level assessment of food insecurity risk due to the pandemic, allowing the government or organizations to allocate more resources or money toward districts that are more affected and allocate fewer resources toward districts that are not affected as much. After this allocation process, the district's local government can target households with higher food insecurity risk as they tend to have more accurate information (World Food Programme, 2015).

This approach is likely better than simple household targeting, especially in the context of developing countries. The process of budgeting for emergency aid based on household poverty status is more challenging and is prone to errors, as the household's status may not be updated annually. More importantly, not all poor households have high food insecurity risk during the pandemic, especially those who work in unaffected industries such as agriculture. In Appendix C, we find that 72.53% of households that were

previously poor in 2019 are not in the high food insecurity risk group during the pandemic. Therefore, using poverty status alone can lead to overbudgeting toward low-risk districts, while our method allows the government and organizations to allocate resources more precisely and thus more effectively. Once the government or international organizations can identify food-insecure households, they can design a suitable food assistance program based on their budget.

It is important to emphasize that our prediction is specifically for short-run food insecurity shocks caused by the pandemic or an emergency and thus should only be used to guide timely responses to provide short-term reliefs during a crisis (Lentz & Barrett, 2013). Such short-term relief policies are very different from social protection programs and policies that aim to address structural causes of food insecurity as they mainly focus on (1) mitigating any negative health effects, especially for children, from short-run shocks (Lentz & Barrett, 2013) and (2) ensuring that families have adequate food during lockdowns or economic downturns so they can follow public health measures.

Intervention options may vary by country, organization, and objective (Lentz, Barrett, Gómez, & Maxwell, 2013; Barrett, Bell, Lentz, & Maxwell, 2009). If budgets are constrained, one option is to send staple food to food-insecure households. Unlike the case of natural disasters, food-insecure households still have access to the market; by reducing the expense of staple food, households can also spend more on non-staple food. Staple foods are also cheaper than non-staple food, which will allow the government to send more to each food-insecure household or reach more food-insecure households.[29] The government and international organizations can also send non-staple foods and micro-nutrient supplements if they have enough money in their budget and enough resources. Although cash transfer is also a popular approach (Gentilini et al., 2020), this relief policy may come with unintended consequences such as higher food prices in remote markets (Filmer, Friedman, Kandpal, & Onishi, 2021).

6. Discussion and conclusion

This study proposes an approach to predict changes in food insecurity risk caused by income shocks at the locality level during emergencies such as the global pandemic. We apply this method to predict changes in food insecurity risk due to the pandemic for all 702 districts in Vietnam. We predict that the share of food-insecure households increases by 0.82 percentage points, on average, as most districts only experience a 0 to 2 percentage point increase, and only 102 out of 702 districts experience increases larger than 2 percentage points. A small number of districts experience an increase as large as 7.86 percentage points. The share of food-insecure children age 0 to 5 is predicted to increase by 0.997 percentage points, although a small number of districts can experience up to 19.33 percentage points.

Several important points can be drawn from these predictions. First, Vietnam is among one of the most successful countries at containing the virus at an early stage, and the overall impact on the economy during the study period was relatively small (Fund, 2020). As a result, our prediction suggests that the average effect on food insecurity across the country might have been quite small during this period. However, we show that a small number of districts are predicted to be affected more severely than others because the pandemic has different impacts on different economic sectors. This finding highlights the importance of effective targeting because these districts may require substantially more support than others. This is especially true for districts with more young children who are affected. We also note that if Vietnam had not contained the pandemic early, food insecurity might have been substantially worse, suggesting effective public health responses to the pandemic can mitigate its potential economic impacts.

The approach proposed in this study and its findings come with several caveats. First, we focus on predicting changes in food insecurity due to income shocks, so our ability to speak about the actual effect of the pandemic on food insecurity is limited. Specifically, our prediction for the effect on food insecurity caused by the pandemic is likely a lower bound because we do not take into account increases in food prices[30] and reductions in remittances and social transfers[31] because of the lack of disaggregated data on changes in food prices and these additional incomes. Higher food prices mean non-agricultural households are more likely to be affected, while lower remittance means households' ability to afford nutritious food decreases even more.[32,33] Second, our IV-based method provides a prediction about changes in food insecurity due to income shocks, so its ability to predict the level of food insecurity is likely insufficient compared to existing prediction methods that incorporate more data such as market price, weather conditions, and geographic characteristics (e.g., Lentz et al., 2019).

This method will nonetheless allow a better targeting approach than most alternative targeting methods given the same constraint of the data. For instance, a common alternative approach is targeting households that are already poor before the pandemic since they are likely more vulnerable to the pandemic income shock. However, the government would also target poor households with low food insecurity risk (errors of inclusion), especially those with household members working in unaffected industries, while missing near-poor households that actually experience food insecurity (errors of exclusion). We assess each type of errors in Appendix C. Targeting previously poor households would include 56.38% poor households that do not have high risk and would miss 11.05% of high-risk households that are not poor. This is highly inefficient, as the amount being sent to the poor but food-secure household can be sent to households that actually need them more.

Another alternative method is to target districts with the largest income reduction due to the pandemic or to target districts with the largest increase in poverty before any household or individual targeting. These methods may also mistarget districts with a smaller increase in food insecurity risk or miss districts that may be experiencing significant increases in food insecurity risk. In Appendix C, we show that 75.96% of districts in the top 10% of income loss are not in the top 10% of (predicted) increase in food insecurity. Similarly, 90.35% of districts in the top 10% of increases in poverty are not in the top 10% of (predicted) increase in food insecurity. Moreover, these two targeting methods would also miss 72.31% and 77.82% of districts in the top 10% of (predicted) increase in food insecurity. These simpler methods are associated with very high error rates, which, again, can be very costly for the government.

[29] For example, the United Nations World Food Programme provides sorghum, corn, millet and rice, beans, black-eyed peas, and vegetable oil in the eight provinces of Chad for three months as a response to temporary food insecurity shocks caused by the pandemic.

[30] Dietrich, Giuffrida, Martorano, and Schmerzeck (2021) show that stringent lockdowns and reduced mobility can increase food prices via higher trade costs, especially in integrated markets in 44 low and middle-income countries. Ruan, Cai, and and Jin (2021) also find that stringent lockdowns increase vegetable prices in China.

[31] The World Bank reports that remittance may decrease by 14% in 2021.

[32] In the case of Vietnam, there have not been any official statistics, to the best of our knowledge, on how remittances are affected by the pandemic; The World Bank (2020) indicates that Vietnam was among the top remittance recipients in 2020. According to the VHLSS, the average share of household income coming from remittances and social transfers is roughly 8.68% to 9.59% during the 2010–2018 period.

[33] Gupta et al. (2021) find that lockdowns in a rural region in India decrease household income substantially, forcing households to reduce their food consumption.

Predictive methods with high accuracy can be used in an effective targeting approach, which in turn will allow preventive measures such as providing food aid to substantially reduce the impacts from food insecurity shocks Barrett (2010). Our method provides a more accurate prediction of food insecurity shocks at the locality level (relative to OLS-based prediction), which the government or international organizations can use to allocate resources and money based on the predicted impacts on food insecurity. A better predicting and targeting approach will allow communities with higher needs to receive more assistance and mitigate any negative impacts of the food insecurity shocks in a timely manner.

CRediT authorship contribution statement

Khoa Vu: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. **Nguyen Dinh Tuan Vuong:** Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. **Tu-Anh Vu-Thanh:** Conceptualization, Resources, Data curation, Writing - review & editing, Validation. **Anh Ngoc Nguyen:** Conceptualization, Resources, Data curation, Writing - review & editing, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors thank Pukitta Chunsuttiwat, Ian Coxhead, Craig Gundersen, Erik Katovich, Hyunseok Kim, Varan Kitayaporn, Saurabh Singhal, Joseph Weinberg, participants of Agricultural Policy (APRG) Conference 2020 and Student Research Colloquium (UW-Madison), and two anonymous reviewers for helpful comments. The authors also thank Trung Xuan Hoang for providing us with the relevant data. All errors in this paper are our own.

Appendix A

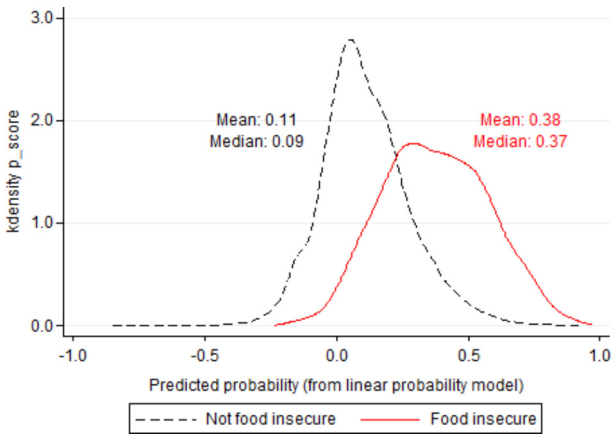


Fig. A.1. Predicted probability distributions by actual food insecurity status. The graph shows the distributions of the predicted probability of food insecurity by household's actual food insecurity status. The predicted probability of food insecurity is obtained from estimating the IV regression on the 2010–2018 VHLSS data. See text for more details.

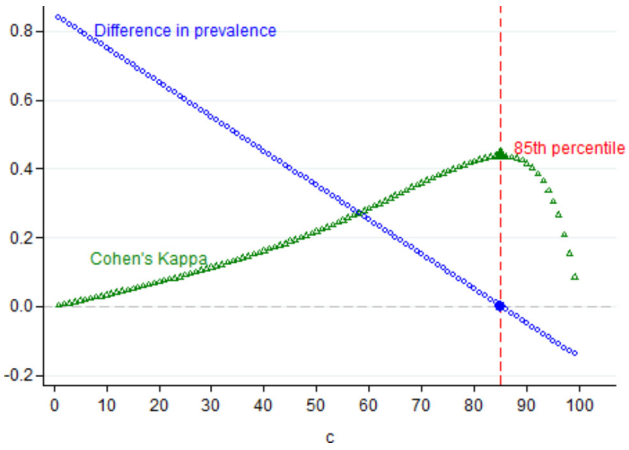


Fig. A.2. Difference in prevalence and Cohen's Kappa statistics to choose the optimal threshold for high-risk households for the linear probability model. For each value c between 0th to 99th percentile, we use c as a threshold to classify the "high-risk" household (based on the predicted probability) and calculate the difference between predicted prevalence, i.e. the share of "high-risk" households minus the share of households that are actually food insecure, and the Cohen's Kappa statistic. We plot the differences in prevalence and the kappa statistics for all thresholds in this graph. The predicted probability is generated from estimating the IV model as explained in Section 3.3.

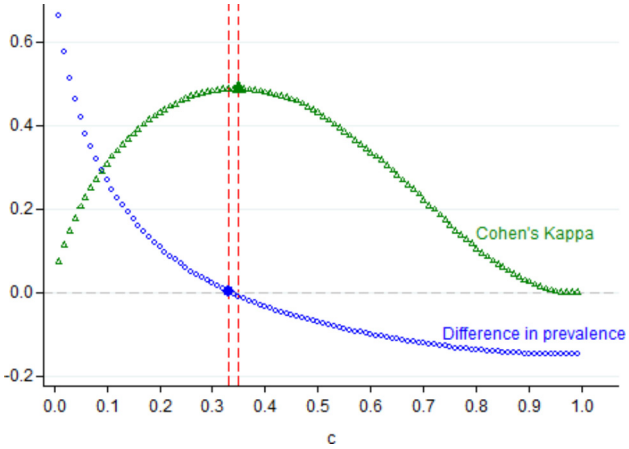


Fig. A.3. Difference in prevalence and Cohen's Kappa statistics to choose the optimal threshold for high-risk households for the IV probit model. For each value c between 0 and 0.99, we use c as a threshold to classify the "high-risk" household (based on the predicted probability) and calculate the difference between predicted prevalence, i.e. the share of "high-risk" households minus the share of households that are actually food insecure, and the Cohen's Kappa statistic. We plot the differences in prevalence and the kappa statistics for all thresholds in this graph. The predicted probability is generated from estimating the IV probit model as explained in Section 3.3. The threshold that minimizes the difference in prevalence is 33%, and the threshold that maximizes the Cohen's Kappa statistic is 35%.

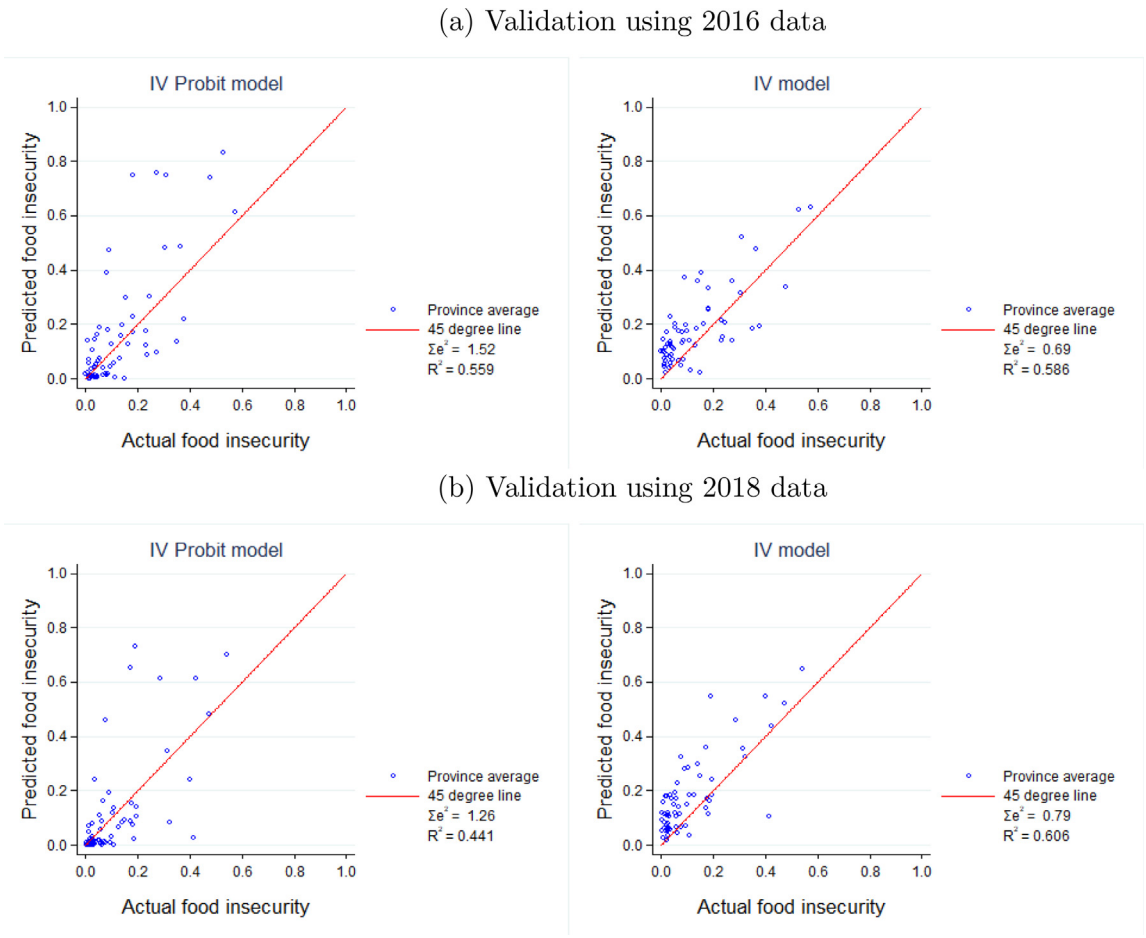


Fig. A.4. Predicted and actual food insecurity of 2018 at the provincial level for IV probit and linear IV. Note: *Predicted food insecurity* measures each province's share of households with high predicted food insecurity risk; *Actual food insecurity* measures the share of households with actual food insecurity in each province using the VHLSS sample in year *y*. Predicted food insecurity risk is obtained by estimating the effect of income on food insecurity using the pre-*y* VHLSS sample and predicting using the LFS sample in *y*.

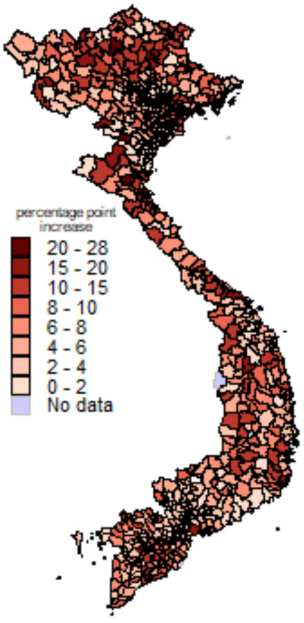


Fig. A.5. District-level percentage point change in the share of households with food insecurity risk before and after COVID-19 under annualized changes. The map shows the percentage point changes in the share of households with high food insecurity risk. Pre-pandemic risk is predicted food insecurity using 2019 household income data from the 2019 Labor Force Survey. Post-pandemic risk is predicted food insecurity using post-pandemic income, which is calculated using the annualized version of the quarterly percentage change of average income by sector from Morisset et al. (2020): decreased by 12.6% in agriculture, 22% in manufacturing, and 32.55% in services per year.

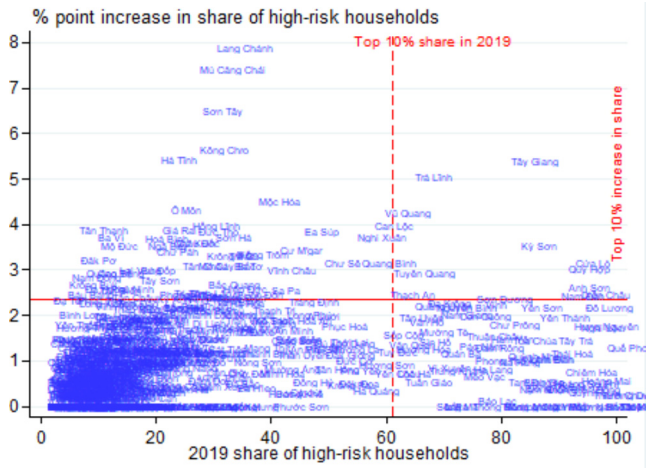


Fig. A.6. Districts' share of households with high food insecurity risk in 2019 and districts' increase in the share of high-risk households due to the pandemic. Note: The horizontal axis shows each district's share of households with high food insecurity risk in 2019. The vertical axis shows districts' increase in the share of high-risk households due to the pandemic. The dashed line shows the top 10% food insecurity in 2019, and the solid line shows the top 10% (predicted) increase in food insecurity due the income shock during the pandemic.

Table A.1
First-stage estimates for income effects on food insecurity.

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Bartik IV						
	1.361***	1.587***	1.567***	0.246***	0.368***	0.440***
	(0.044)	(0.047)	(0.063)	(0.068)	(0.071)	(0.075)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
N	46443	46443	27462	46443	46443	46443
Panel B: Non-ag. Bartik IV						
	1.444***	1.669***	1.902***	0.189***	0.472***	0.421***
	(0.050)	(0.055)	(0.074)	(0.072)	(0.078)	(0.081)
	[0.000]	[0.000]	[0.000]	[0.009]	[0.000]	[0.000]
N	46443	46443	27462	46443	46443	46443
Additional controls						
Province FE	✓			✓	✓	✓
District FE		✓				
Household FE			✓			
Year FE				✓		✓
Province FE x linear trends					✓	
Province x Year FE						✓

The table reports the first-stage results from the IV estimations in Tables 2 and 6. Standard errors are clustered at the district level and are reported in parentheses. P-values are reported in brackets. The household fixed effects model does not include 2009 controls. All models control for urban, household size, and the fraction of households with a postsecondary education. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix B. Validity of food insecurity measurement

In this section, we discuss the validity of the paper's main measure of food insecurity. We compare the main measure of food insecurity of this study, that is, when a household's staple calorie share is above 84%, with two other food insecurity measures that are commonly used: the measure using the HDDS and the self-reported food insecurity. Given that food insecurity is strongly related to the ability to afford everyday meals, a valid food insecurity measure would strongly correlate with household income and wealth. One may also expect that rural households and ethnic minority households would be more likely to be food insecure as they are more disadvantaged than urban and Kinh households. In Fig. B.1, we plot the percentage of households classified as food



Fig. B.1. Food-insecure household by household income and wealth deciles. The figure reports the share of food-insecure households by income deciles, wealth index deciles, and ethnicity. A lower decile means lower income or lower wealth. The wealth index is constructed using principal component analysis for electricity, piped water, air conditioner, computer, washing machine, refrigerator, television, and radio.

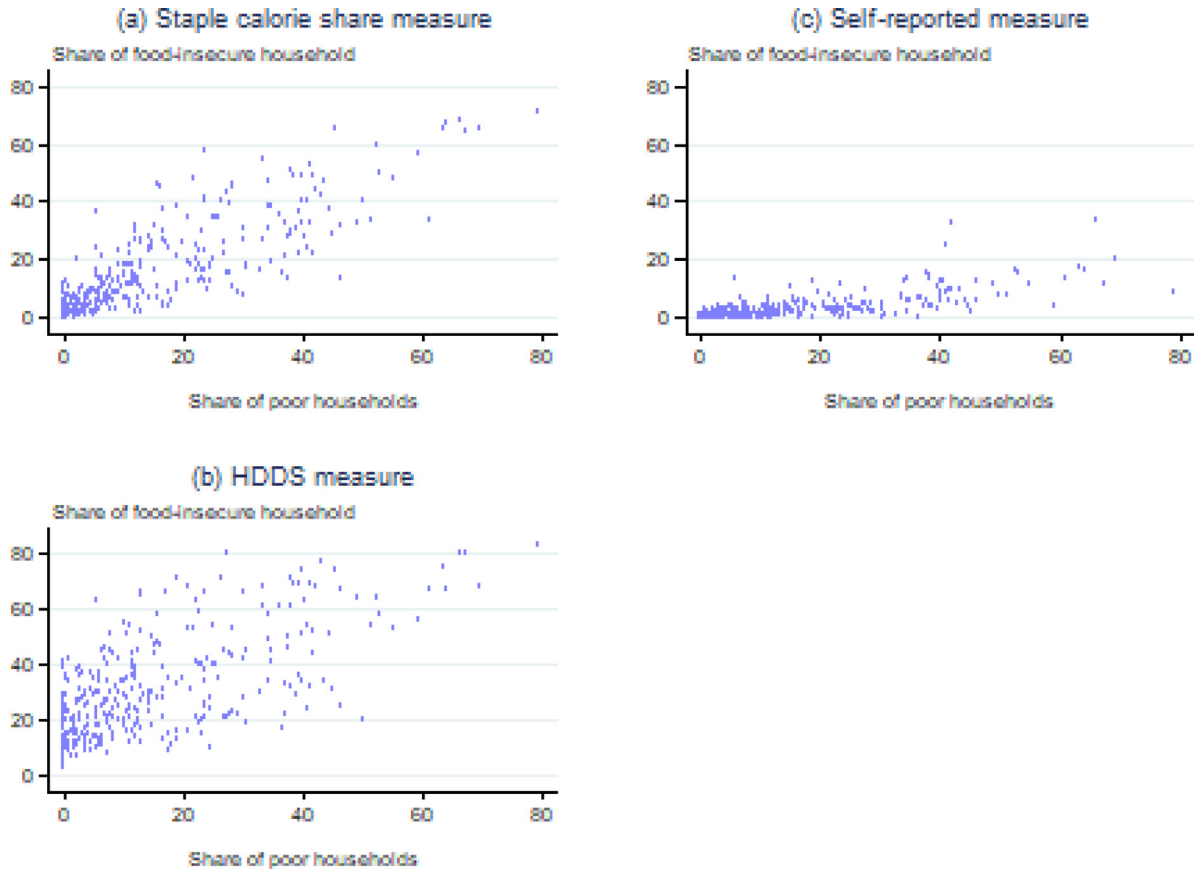


Fig. B.2. Food insecurity and poverty by province and year. The figure shows scatter plots for shares of food-insecure households and shares of poor households at the province-year level for the 2010–2018 data. Each dot is a province-year observation. Figure (a) shows food insecurity using the staple calorie share approach, (b) shows food insecurity using the HDDS approach, and (c) shows self-reported food insecurity.

insecure using different measures by household income deciles (Figure (a)) and household wealth deciles (Figure (b)) for urban and rural households separately. In Figure (c), we also plot the percentage of food-insecure households by the household head's ethnicity.

All three measures of food insecurity decrease as household income and wealth increase, and this is true for both rural and urban households. However, self-reported food insecurity is considerably lower than the other two measures. This discrepancy is partly due to selection bias: households only answer the question about food insecurity if they have an official poverty status, so poor households and near-poor households without such a status would not be considered as food insecure. Another reason is that the staple calorie share approach and the HDDS approach focus on the quality of diet besides quantity, so food-insecure households may have a similar number of meals but a very different quality than food-secure households. Our staple calorie share approach is more similar to the HDDS approach for this reason as well. We also find that all three measures of food insecurity are higher among ethnic minority households than Kinh households.

These patterns are also consistent when we aggregate the data at the province level. In Fig. B.2, we graph the scatter plots of different food insecurity measures against poverty at the province level, which tells a similar story—food insecurity measured by the staple calorie share and by HDDS is strongly and positively correlated with poverty, while self-reported food insecurity is positively correlated with poverty but not as strongly.

These results suggest that the food insecurity measure using the staple calorie share is valid since it is strongly correlated with income, wealth, and poverty, similar to the other common measures using the HDDS and self-reported data. The staple calorie share approach is better than the HDDS approach because it uses a theoretically derived cutoff for food insecurity, and it is better than the self-reported measure because it does not suffer from the selection bias described above.

Appendix C. Comparison between different targeting approaches

An important question to ask is whether using the targeting method proposed in this study is much better than using simpler methods such as targeting poor households, targeting districts that suffer the largest income reductions, or targeting districts with the largest increases in poverty. In this section, we calculate the inclusion error rate (IER) and exclusion error rate (EER) for each of these three approaches to measure the extent to which these alternative methods include wrong households/districts and exclude right households/districts (Brown et al., 2018).

Consider the poor household targeting approach: the government sends food aid to households below the national poverty line *before* the pandemic. Let $poor_i^{2019}$ be a binary variable that indicates whether household i is poor in 2019, and let $\widehat{insecurity}_i$ be a binary variable that indicates whether the household has high food insecurity risk using the prediction approach of this study.[34] Let w_i denote the household's sampling weight from the LFS. The IER is

[34] Recall that food insecurity risk is the predicted value from Eq. 2 using the post-pandemic income, and high-risk households are those with risk above the 85th percentile cutoff.

the proportion of poor, but food-secure households are targeted and can be calculated as

$$IER = \frac{\sum_{i}^{N} w_i.1\left(poor_i = 1 | \widehat{insecurity}_i = 1\right)}{\sum_{i}^{N} w_i.1(poor_i = 1)}.$$

In contrast, the EER is the proportion of food-insecure households that are not targeted because they are not poor before the pandemic, and it is calculated as

$$EER = \frac{\sum_{i}^{N} w_i.1\left(\widehat{insecurity}_i = 1 | poor_i = 0\right)}{\sum_{i}^{N} w_i.1\left(\widehat{insecurity}_i = 1\right)}.$$

The IER of the poor household targeting approach is 56.38% (95% CI: 55.84, 56.93), and the EER is 11.05% (95% CI: 10.46, 11.63). In other words, 56.38% of targeted households are actually not food insecure. Only 11.05% of food-insecure households would be missed using this method.

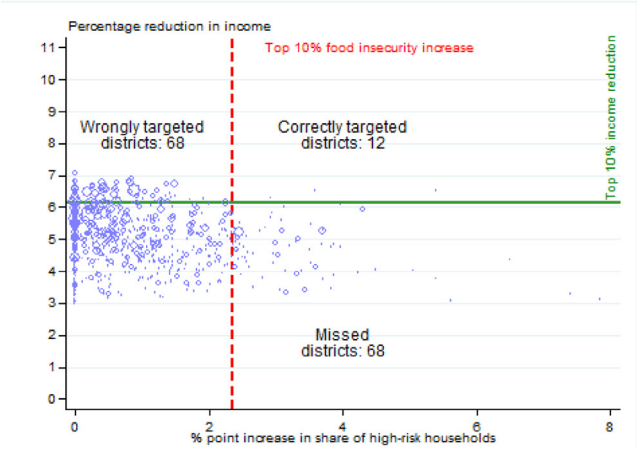
Next, we use the same approach to assess two other alternative approaches: targeting districts with the highest income losses and targeting districts with the largest poverty increases. To do this, we estimate changes in income and poverty due to COVID-19 for each district using the same method to predict changes in food insecurity. Specifically, we use the 2019 LFS household data and the industry-specific income shocks to calculate the new income for each household due to these shocks, and then we estimate the percentage change in income for each district. We also use the original income and new income to identify the poverty status, that is, monthly income per capita below the national poverty line, for each household before and after the income shock. We then estimate the percentage point change in poverty for each district.

Fig. C.1 shows the scatter plots of district-level changes in food insecurity (our main result) against changes in income and changes in poverty. To see why using income targeting and poverty targeting may lead to ineffective targeting, suppose the government chooses to target districts in the top 10% of those that experienced an increase in food insecurity; that is, their predicted food insecurity increases are in the 90th percentile. In other words, the government would want to target all districts to the right of the red dashed line in the plots.

Now suppose that the government does not use food insecurity targeting, but instead target districts in the top 10% of those that saw a reduction in income; that is, their predicted income reductions are in the 90th percentile. In other words, the government would target all districts above the green solid line in Figure (a). In this case, all districts to the left of the red dashed line and above the green solid line would be wrongly targeted as they experience a food insecurity increase that is below the 90th percentile. In contrast, all districts to the right of the red dashed line and below the green solid line would be missed because they experience food insecurity increase in the 90th percentile but would not be targeted. The correctly targeted districts are the dots in the first quadrant of the graph. In this scenario, the government would wrongly target 68 districts, miss 68 districts, and correctly target 12 districts. We apply the same IER and EER formulas for districts instead of households and calculate each district's as one over the district's population. The IER is 75.96% (95% CI: 66.80, 85.13) and the EER is 72.31% (95% CI: 62.37, 82.25).

Now suppose that the government targets districts in the top 10% of districts that experienced an increase in poverty; that is, their predicted poverty increases are in the 90th percentile—all districts above the green solid line in Figure (b). Similarly, this targeting would mistarget all districts to the left of the red dashed line and above the green solid line but would miss all districts to the right of the red line and below the green line. In this scenario, the government would wrongly target 59 districts, miss 20 districts, and correctly target 60 districts. The IER here would be 90.35% (95% CI: 85.64, 95.05) and the EER 77.82% (95% CI: 68.64, 87.00).

(a) Food insecurity targeting versus income targeting



(b) Food insecurity targeting versus poverty targeting

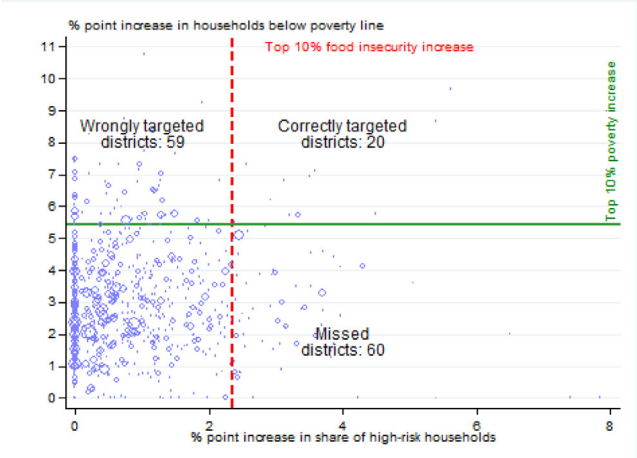


Fig. C.1. Food insecurity change targeting versus income change targeting and poverty change targeting. Note: Figure (a) shows a scatter plot of the predicted percentage reduction in income against the predicted percentage point increase in food insecurity. Figure (b) shows a scatter plot of the predicted percentage point increase in poverty against the predicted percentage point increase in food insecurity. Each dot represents a district. The red dashed line shows the 90th percentile of the increase in food insecurity. The green solid line shows the 90th percentile of the income reduction in Figure (a) and the 90th percentile of poverty increase in Figure (b).

These simple exercises suggest that although targeting using pre-pandemic poverty statuses of households or using income losses is simpler, the error rates associated with them are very high; using them means the government would spend a significant amount of resources on districts it does not intend to and miss the districts that actually need those resources.

Appendix D. Comparison between IV-based prediction and 2020 data

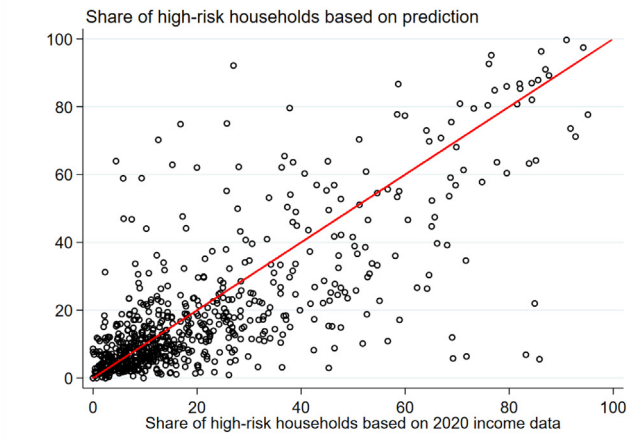
A major concern of our predictive approach is that we exploit local exposure to industry-wide labor demand shocks to household incomes to identify the effect of income on food insecurity. We also

rely on external information about sector-wide income shocks and assume that all workers within the same sector have an equal chance to receive such shocks. The income shocks from the pandemic might be very different in nature. If the pandemic income shocks are driven by unemployment in a nonrandom way, that is, more vulnerable workers lose their jobs and income, then the IV-based prediction is no longer accurate.

We address this concern by comparing our IV-based prediction with the food insecurity measure generated from 2020 household income data (using the 2020 LFS data). If there is a large difference between our prediction and the food insecurity measure that generated the actual data, we can conclude that the income shocks from our IV estimation are not very compatible to make predictions about income shocks caused by a pandemic.

We plot each district's share of high-risk households based on 2020 income data and the share of high-risk households based on our prediction in Fig. D.1 (a) along with a 45-degree line as the benchmark. In Figure (b), we plot the distribution of the percentage point difference (in absolute value) between the two variables. We observe in Figure (a) that our IV-based prediction is fairly consistent with the food insecurity measure generated from the actual 2020 income data, although we also observe a few outlier districts. In Figure (b), we find that the difference between the two variables is 10 percentage points or less for 82.18% of districts. In other words, our predicted food insecurity measure is fairly close to the food insecurity measure based on the actual income data in 2020. This finding suggests our approach is compatible to predict food insecurity during a pandemic, even though the income shocks caused by the pandemic might be driven mainly through unemployment or furlough.

(a) Prediction vs. reality



(b) The distribution of the prediction-reality difference

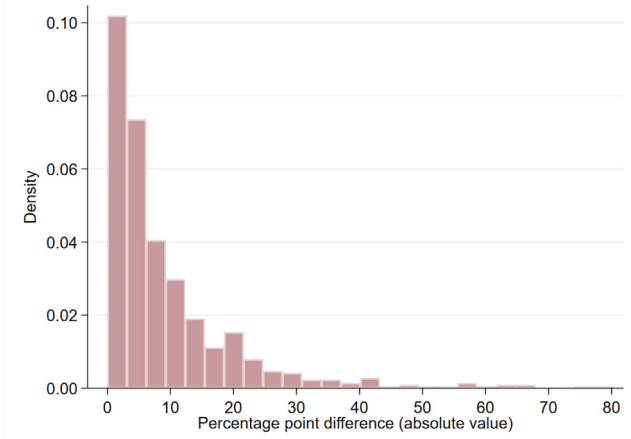


Fig. D.1. Differences in share of "high-risk" households based on prediction and share of "high-risk" households based on 2020 income. Note: Figure (a) shows districts' share of high-risk households based on 2020 income data and districts' shares of high-risk households based on our prediction. Figure (b) shows the distribution of differences between these two shares.

References

Abadie, A., Athey, S., Imbens, G.W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? National Bureau of Economic Research..

Abebaw, D., Admassie, A., Kassa, H., & Padoch, C. (2020). Can rural outmigration improve household food security? Empirical evidence from Ethiopia. *World Development, 129* 104879.

Aggarwal, S., Jeong, D., Kumar, N., Park, D.S., Robinson, J., and Spearot, A. (2020). Did covid-19 market disruptions disrupt food security? evidence from households in rural liberia and malawi. Technical report, National Bureau of Economic Research..

Ahn, S., & Norwood, F. B. (2020). Measuring food insecurity during the covid-19 pandemic of spring 2020. *Applied Economic Perspectives and Policy*.

Aiken, E., Bellue, S., Karlan, D., Udry, C. R., & Blumenstock, J. (2021). *Machine learning and mobile phone data can improve the targeting of humanitarian assistance* Technical report. National Bureau of Economic Research.

Amare, M., Abay, K. A., Tiberti, L., & Chamberlin, J. (2021). Covid-19 and food security: Panel data evidence from nigeria. *Food Policy, 101* 102099.

Andree, B.P.J., Chamorro, A., Kraay, A., Spencer, P., and Wang, D. (2020). Predicting food crises..

Athey, S. (2018). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda* (pp. 507–547). University of Chicago Press.

Auffhammer, M., & Steinhauser, R. (2012). Forecasting the path of us co2 emissions using state-level information. *Review of Economics and Statistics, 94*(1), 172–185.

Barrett, C. B. (2010). Measuring food insecurity. *Science, 327*(5967), 825–828.

Barrett, C. B., Bell, R., Lentz, E. C., & Maxwell, D. G. (2009). Market information and food insecurity response analysis. *Food Security, 1*(2), 151–168.

Barrett, C.B., Upton, J., Tennant, E., and Florella, K. (2021). Household resilience, and rural food systems: Evidence from southern and eastern africa. Available at SSRN 3992671..

Baylis, K., Fan, L., & Nogueira, L. (2019). Agricultural market liberalization and household food insecurity in rural China. *American Journal of Agricultural Economics, 101*(1), 250–269.

BBC (2020). Coronavirus: Vietnamese farmers cry as exports stuck at border. https://www.bbc.com/vietnamese/vietnam-51353578. [Online; accessed 4th May 2021]..

BBC (2021). Vietnam: Increasing risk of a new wave of covid-19 infection? https://www.bbc.com/vietnamese/vietnam-56939029. [Online; accessed 4th May 2021]..

Ben-David, A. (2007). A lot of randomness is hiding in accuracy. *Engineering Applications of Artificial Intelligence, 20*(7), 875–885.

Ben-David, A. (2008). About the relationship between roc curves and cohen's kappa. *Engineering Applications of Artificial Intelligence, 21*(6), 874–882.

Béné, C., Bakker, D., Chavarro, M. J., Even, B., Melo, J., & Sonneveld, A. (2021). Global assessment of the impacts of covid-19 on food security. *Global Food Security, 31* 100575.

Bickel, G., Nord, M., Price, C., Hamilton, W., and Cook, J. (2000). Guide to measuring household food security, Revised 2000,. U.S. Department of Agriculture, Food and Nutrition Service,, Alexandria VA..

Bitler, M., Hoynes, H.W., and Schanzenbach, D.W. (2020). The social safety net in the wake of COVID-19. National Bureau of Economic Research..

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science, 350*(6264), 1073–1076.

Brown, C., Ravallion, M., & van de Walle, D. (2018). A poor means test? econometric targeting in africa. *Journal of Development Economics, 134*(C), 109–124.

Brown, C.S., Ravallion, M., and Van De Walle, D. (2020). Can the World's Poor Protect Themselves from the New Coronavirus? National Bureau of Economic Research..

Chaaban, J., Ghattas, H., Irani, A., & Thomas, A. (2018). Targeting mechanisms for cash transfers using regional aggregates. *Food security, 10*(2), 457–472.

Christian, P., Kandpal, E., Palaniswamy, N., & Rao, V. (2019). Safety nets and natural disaster mitigation: evidence from cyclone phailin in odisha. *Climatic Change, 153*(1), 141–164.

Dang, H.-A. and Giang, L. (2020). Turning vietnam's covid-19 success into economic recovery: A job-focused analysis of individual assessments on their finance and the economy..

Deitchler, M., Ballard, T., Swindale, A., and Coates, J. (2010). Validation of a measure of household hunger for cross-cultural use: Food and nutrition technical assistance ii project (fanta-2). washington, dc: Fanta-2..

Devereux, S., Béné, C., & Hoddinott, J. (2020). Conceptualising covid-19's impacts on household food security. *Food Security, 12*(4), 769–772.

Dietrich, S., Giuffrida, V., Martorano, B., & Schmerzeck, G. (2021). Covid-19 policy responses, mobility, and food prices. *American Journal of Agricultural Economics*.

Filmer, D., Friedman, J., Kandpal, E., & Onishi, J. (2021). Cash transfers, food prices, and nutrition impacts on ineligible children. *The Review of Economics and Statistics*, 1–45.

Freeman, E. A., & Moisen, G. G. (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological modelling, 217*(1–2), 48–58.

Gatto, M., & Islam, A. H. M. S. (2021). Impacts of covid-19 on rural livelihoods in bangladesh: Evidence using panel data. *PloS one, 16*(11) e0259264.

Gelli, A., Aberman, N.-L., Margolies, A., Santacroce, M., Baulch, B., & Chirwa, E. (2017). Lean-season food transfers affect children's diets and household food security: evidence from a quasi-experiment in malawi. *The Journal of Nutrition, 147*(5), 869–878.

Gentilini, U., Almenfi, M., Orton, I., and Dale, P. (2020). Social protection and jobs responses to covid-19..

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review* (forthcoming).

GSO (2020a). Report on the impact of the COVID-19 on labour and eployment situation in Vietnam. Accessed on Sep 1st 2020 via https://gso.gov.vn/default_en.aspx?tabid=462&idmid=2&ItemID=19678..

GSO (2020b). Báo cáo Kết quả Khảo sát đánh giá tác động của dịch COVID-19 đến hoạt động sản xuất, kinh doanh của doanh nghiệp [Firm survey summary report on the impact of COVID-19]. Accessed on Sep 1st 2020 via https://www.gso.gov.vn/default.aspx?tabid=382&idmid=2&ItemID=19623..

GSO (2021). Report on the impact of the COVID-19 Pandfemic on labour and eployment situation of the 2nd Quarter of 2021. Accessed on July 7th 2021 via https://www.gso.gov.vn/en/data-and-statistics/2021/07/report-on-impact-of-covid-19-pandemic-on-labour-and-employment-of-the-second-quarter-of-2021/..

Gundersen, C., Hake, M., Dewey, A., & Engelhard, E. (2020). Food insecurity during covid-19. *Applied Economic Perspectives and Policy*.

Gundersen, C., Kreider, B., & Pepper, J. (2011). The economics of food insecurity in the united states. *Applied Economic Perspectives and Policy, 33*(3), 281–303.

Gundersen, C., Kreider, B., & Pepper, J. V. (2017). Partial identification methods for evaluating food assistance programs: A case study of the causal impact of SNAP on food insecurity. *American Journal of Agricultural Economics, 99*(4), 875–893.

Gupta, A., Zhu, H., Doan, M. K., Michuda, A., & Majumder, B. (2021). Economic impacts of the covid- 19 lockdown in a remittance-dependent region. *American Journal of Agricultural Economics, 103*(2), 466–485.

Hidrobo, M., Hoddinott, J., Kumar, N., & Olivier, M. (2018). Social protection, food security, and asset formation. *World Development, 101*, 88–103.

Hirvonen, K., de Brauw, A., & Abate, G. T. (2021). Food consumption and food security during the covid-19 pandemic in addis ababa. *American Journal of Agricultural Economics, 103*(3), 772–789.

International Monetary Fund (2020). World economic outlook, Oct 2020: A long and difficult ascent. International Monetary Fund..

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (volume 112). Springer.

Jensen, R.T. and Miller, N.H. (2010). A revealed preference approach to measuring hunger and undernutrition. NBER Working Paper No. 16555..

Jones, A. D., Shrinivas, A., & Bezner-Kerr, R. (2014). Farm production diversity is associated with greater household dietary diversity in malawi: Findings from nationally representative data. *Food Policy, 46*, 1–12.

Kuku, O., Gundersen, C., & Garasky, S. (2011). Differences in food insecurity between adults and children in zimbabwe. *Food Policy, 36*(2), 311–317.

Kuma, T., Dereje, M., Hirvonen, K., & Minten, B. (2019). Cash crops and food security: Evidence from ethiopian smallholder coffee producers. *The Journal of Development Studies, 55*(6), 1267–1284.

Lentz, E., Michelson, H., Baylis, K., & Zhou, Y. (2019). A data-driven approach improves food insecurity crisis prediction. *World Development, 122*, 399–409.

Lentz, E. C., & Barrett, C. B. (2013). The economics and nutritional impacts of food assistance policies and programs. *Food Policy, 42*, 151–163.

Lentz, E. C., Barrett, C. B., Gómez, M. I., & Maxwell, D. G. (2013). On the choice and impacts of innovative international food assistance instruments. *World Development, 49*, 1–8.

Lewbel, A., Dong, Y., & Yang, T. T. (2012). Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics/Revue canadienne d'économique, 45*(3), 809–829.

Maxwell, D., Vaitla, B., & Coates, J. (2014). How do indicators of household food insecurity measure up? an empirical comparison from ethiopia. *Food Policy, 47*, 107–116.

Ministry of Health (2021). New information from the ministry of health about the third wave. https://ncov.moh.gov.vn/web/guest/-/6847426-2117. [Online; accessed 4th May 2021]..

Ministry of Industry and Trade (2020). Báo cáo tình hình sản xuất công nghiệp và hoạt động thương mại tháng 7 và 7 tháng đầu năm 2020, giải pháp thực hiện trong thời gian tới [Report on industrial production and trade in July and the first 7 months of 2020: Policy implication]. Accessed on Sep 1st 2020 via https://www.moit.gov.vn/web/guest/bao-cao-tong-hop1..

Minnesota Population Center (2019). Integrated public use microdata series, International: Version 7.2 [dataset]. IPUMS, Minneapolis, MN. https://doi.org/10.18128/D020.V7.2..

Mishra, K., & Rampal, J. (2020). The covid-19 pandemic and food insecurity: A viewpoint on india. *World Development, 135* 105068.

Morisset, J., Dinh, V.T., Doan, Q.H., Pham, D.M., Mandani, D., Pimhidzai, O., Kaiser, K.-A., Do, D.V., Alatabani, A.F., and Yang, J. (2020). Taking stock: What will be the new normal for Vietnam? - The Economic impact of COVID-19. Accessed via http://documents.worldbank.org/curated/en/101991595365511590/Taking-Stock-What-will-be-the-New-Normal-for-Vietnam-The-Economic-Impact-of-COVID-19..

National Institute of Nutrition (2007). *Vietnamese food composition table*. Hanoi, Vietnam: Ministry of Health.

Newell, R.G., Prest, B.C., and Sexton, S. (2018). The gdp-temperature relationship: implications for climate change damages. Resour. Future Work. Pap..

Nguyen, P. H., Kachwaha, S., Pant, A., Tran, L. M., Ghosh, S., Sharma, P. K., Shastri, V. D., Escobar-Alegria, J., Avula, R., & Menon, P. (2021). Impact of covid-19 on household food insecurity and interlinkages with child feeding practices and coping strategies in uttar pradesh, india: a longitudinal community-based study. *BMJ open, 11*(4) e048738.

Nikulkov, A., Barrett, C. B., Mude, A. G., & Wein, L. M. (2016). Assessing the impact of us food assistance delivery policies on child mortality in northern kenya. *PloS one, 11*(12) e0168432.

Niles, M. T., & Brown, M. E. (2017). A multi-country assessment of factors related to smallholder food security in varying rainfall conditions. *Scientific Reports, 7*(1), 1–11.

Paslakis, G., Dimitropoulos, G., & Katzman, D. K. (2020). A call to action to address covid-19–induced global food insecurity to prevent hunger, malnutrition, and eating pathology. *Nutrition reviews*.

Picchioni, F., Goulao, L. F., & Roberfroid, D. (2021). The impact of covid-19 on diet quality, food security and nutrition in low and middle income countries: A systematic review of the evidence. *Clinical Nutrition*.

Ratcliffe, C., McKernan, S.-M., & Zhang, S. (2011). How much does the supplemental nutrition assistance program reduce food insecurity? *American Journal of Agricultural Economics, 93*(4), 1082–1098.

Reardon, T., Bellemare, M. F., Zilberman, D., et al. (2020). How covid-19 may disrupt food supply chains in developing countries. *IFPRI Book Chapters*, 78–80.

Ruan, J., Cai, Q., and Jin, S. (2021). Impact of covid-19 and nationwide lockdowns on vegetable prices: Evidence from wholesale markets in china. American journal of agricultural economics..

Savy, M., Fortin, S., Kameli, Y., Renault, S., Couderc, C., Gamli, A., Amouzou, K., Perenze, M., & Martin-Prével, Y. (2020). Impact of a food voucher program in alleviating household food insecurity in two cities in senegal during a food price crisis. *Food Security*, 1–14.

Schanzenbach, D., & Pitts, A. (2020). How much has food insecurity risen? evidence from the census household pulse survey. *Institute for Policy Research Rapid Research Report*.

Swindale, A. and Bilinsky, P. (2006). Household dietary diversity score (hdds) for measurement of household food access: Indicator guide..

Swindale, A. and Bilinsky, P. (2016). Household dietary diversity score (HDDS) for measurement of household food access: Indicator guide (v.2). FHI 360/FANTA, Washington, D.C..

The World Bank (2018). Climbing the ladder: Poverty reduction and shared prosperity in Vietnam..

The World Bank (2020). In Global outlook: Pandemic, recession: The global economy in crisis..

The World Bank (2020). Migration and development brief 33 October 2020–Phase II: COVID-19 Crisis through a Migration Lens..

The World Bank (2020). Vietnam macro monitoring (english). Accessed via http://documents1.worldbank.org/curated/en/213631600443823992/pdf/Vietnam-Macro-Monitoring.pdf..

Thi, H. T., Simioni, M., & Thomas-Agnan, C. (2018). Assessing the nonlinearity of the calorie-income relationship: An estimation strategy - With new insights on nutritional transition in Vietnam. *World Development, 110*, 192–204.

Tuoi Tre online (2020). Stop all flights between vietnam and china starting from this evening. https://tuoitre.vn/dung-toan-bo-chuyen-bay-giua-viet-nam-va-trung-quoc-tu-chieu-nay-20200201155919672.htm. [Online; accessed 4th May 2021]..

Vaitla, B., Coates, J., Glaeser, L., Hillbruner, C., Biswal, P., Maxwell, D., et al. (2017). The measurement of household food security: Correlation and latent variable analysis of alternative indicators in a large multi-country dataset. *Food Policy, 68* (C), 193–205.

Vaitla, B., Coates, J., & Maxwell, D. (2015). *Comparing household food consumption indicators to inform acute food insecurity phase classification* (p. 360). Washington, DC: FHI.

Verpoorten, M., Arora, A., Stoop, N., & Swinnen, J. (2013). Self-reported food insecurity in africa during the food price crisis. *Food Policy, 39*, 51–63.

Vietnam National Administration of Tourism (2021a). International tourism revenue. https://vietnamtourism.gov.vn/index.php/statistic/receipts..

Vietnam National Administration of Tourism (2021b). International tourism statistics. http://www.vietnamtourism.gov.vn/index.php/statistic/international?txtkey=&year=2019&period=t12..

Vietnambiz (2020a). Pork prices in vietnam. https://vietnambiz.vn/gia-thit-heo-hom-nay-15-8-giam-10-tai-cua-hang-vissan-20200815084310709.html. [Online; accessed 16th August 2020]..

Vietnambiz (2020b). Rice prices in vietnam. https://vietnambiz.vn/gia-gao.html. [Online; accessed 16th August 2020]..

von Carnap, T. (2021). Remotely-sensed market activity as a short-run economic indicator in rural areas of developing countries. Available at SSRN 3980969..

Webb, P., Coates, J., Frongillo, E. A., Rogers, B. L., Swindale, A., & Bilinsky, P. (2006). Measuring household food insecurity: why it's so important and yet so difficult to do. *The Journal of Nutrition, 136*(5), 1404S–1408S.

Wolfson, J. A., & Leung, C. W. (2020). Food insecurity and covid-19: Disparities in early effects for us adults. *Nutrients, 12*(6), 1648.

World Food Programme (2015). *Community-based targeting guide*. Kenya: World Food Programme.
Yang, J., Panagoulias, P., and Demarchi, G. (2020). Monitoring covid-19 impacts on households in vietnam, report no. 1..
Zhou, Y., Lentz, E., Michelson, H., Kim, C., and Baylis, K. (2021). Machine learning for food security: Principles for transparency and usability. Applied Economic Perspectives and Policy..