Drought Shocks and Firm Performance: A Study of Causal Effects Using Machine Learning

LIJIN LIU[a] AND YILIN WU[a,b]

[a] School of Statistics, Renmin University of China, Beijing, China
[b] Center for Applied Statistics, Renmin University of China, Beijing, China

(Manuscript received 11 October 2024, in final form 9 May 2025, accepted 21 July 2025)

ABSTRACT: This paper explores the effects of drought shocks on firms' performance based on the generalized random forest by integrating farmers and firms into the same analytical framework. The results indicate that drought shocks have a negative impact on firms' revenues, employment size, and profit. The negative impact of drought shocks on firms has significant heterogeneous effects across firms with different assets, industries, and ages. Robustness tests reveal that the data used in the study contain sufficient information so that the results are not affected by omitted variables and other confounding factors. In addition, the results remain robust when nonrandomization, repeated shocks, and spillover effects are considered. Mechanistic analyses reveal that drought shocks affect firms through demand and cost effects. On the one hand, drought shocks affect the consumption level of regional farmers, which reduces the demand for firms' products, thus affecting firms' revenues and employment. On the other hand, drought shocks lead to a decline in demand for firms' products that is greater than the decline in firms' costs, which in turn has an impact on firms' profits. The results of the adaptive strategies analysis show that promoting agricultural technology, improving agricultural facilities, and promoting financial inclusion in the region can effectively mitigate the negative impacts of drought shocks on firms' performance.

KEYWORDS: Social Science; Drought; Extreme events; Climate change; Statistics; Adaptation

1. Introduction

Countries worldwide have been experiencing varying degrees of climate change in recent years, with rising temperatures caused by greenhouse gas emissions being of particular concern (IPCC 2022). Global warming affects the water cycle and alters atmospheric circulation patterns, exacerbating droughts on a global scale (Feng et al. 2020; Straffelini and Tarolli 2023). Countries in Europe, North Africa, and South America have suffered several heat waves in recent years, which have adversely affected various industries. Among these, agriculture is the sector most affected by drought, which has led to a reduction in food production and thus affects food security (Jarrett and Tackie 2024). Besides, the impact of the drought on the agricultural sector affects the secondary and tertiary sectors through the industrial chain, with serious adverse effects on economic development (Grabrucker and Grimm 2021). According to the FAO (2023), over the past 30 years, agriculture has accounted for an average of 23% of all losses caused by climate change–induced extreme weather events across all sectors; more than 65% of the losses suffered by the agricultural sector were caused by droughts, which resulted in losses amounting to about \$3.8 trillion. It is also the case that China has experienced record-high temperatures in recent years, with droughts becoming more frequent, widespread, and intense in all regions. In 2023, the average temperature in China continued to set a new record extreme value, which was the highest in history, and the precipitation was the second lowest since 2012. It is foreseeable that in the future, as the global temperature rises, dry weather events in China will become more frequent and intense, which will adversely affect economic and social development.

Existing research has pointed out that the increasing frequency of drought shocks has affected different production sectors to varying degrees. For example, a study by Amare et al. (2021) reveals that drought shocks lead to an increase in the incidence of pests and diseases and a decline in crop yields. This decrease in crop yields significantly affects farm household incomes and leads to persistent poverty, with the negative impacts greater in developing countries and among asset-poor households (Bozzola and Smale 2020; Branco and Féres 2021). Meanwhile, increasingly frequent drought shocks forced farmers to downsize crop production and impede interfarm land transfers, leading to regional abandonment of arable land and reduced food production, which negatively affects national food security (Chavas et al. 2022). As a result of the reduced supply of products from the agricultural sector, the input costs of the processing industries associated with it increase, affecting the production of enterprises. Reduced enterprise production and higher product prices, in turn, negatively impact household welfare (Grabrucker and Grimm 2021). Besides, drought shocks also affect the labor market significantly. It was found that part-time farmers will give more labor inputs to their crops to stabilize production in the event of a drought shock. This will lead to a reduction in the market labor force, and the resulting losses are estimated to exceed 0.20% of GDP per year (He et al. 2022). Overall,

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/WCAS-D-24-0111.s1.

Corresponding author: Lijin Liu, liulijin2938@163.com

DOI: 10.1175/WCAS-D-24-0111.1

there is a consensus among existing studies that drought shocks harm microsubjects (David et al. 2021; Agamile et al. 2021; Farzana et al. 2022).

While the findings of studies on the impact of drought shocks on the agricultural sector are relatively consistent, studies on the impact of drought shocks on firms' performance have been less addressed and the conclusions obtained are inconsistent (Lazzaroni and van Bergeijk 2014; Clò et al. 2024). On the one hand, neoclassical growth models based on exogenous technological progress in some studies predict more rapid capital accumulation following capital destruction caused by natural disasters (Noy and Vu 2010), which promotes firm growth (Kirchberger 2017). This is reflected in higher growth rates, which are temporarily maintained until a new steady-state equilibrium growth is reached. On the other hand, studies based on endogenous growth models with increasing returns to scale of production have found that disasters reduce the capital stock of firms and have a negative impact on firm performance (Zhou and Botzen 2021). The inconsistent results call for researchers to explore the impact of climate change on firm performance in more detail, to better serve economic development. However, there is still little discussion on how drought shocks affect enterprise performance against the backdrop of the increasing intensity and frequency of drought in China, which has hindered the process of effective response to climate change by the government and the enterprise sector in China. Therefore, it is not only a practical but also a theoretical requirement to further clarify the impact of drought shocks on enterprise performance, investigate the transmission mechanisms involved, and actively explore coping strategies.

Previous research has explored the impacts of natural disasters on the agricultural sector or firm individually, either based on the difference-in-difference or instrumental variables (Leiter et al. 2009; Clò et al. 2024). However, the farm sector and the firm sector do not exist independently of each other without any linkage, and they are inextricably linked (Grabrucker and Grimm 2021). Therefore, exploring the impacts of drought shocks on these two sectors separately does not provide a complete picture of the impacts of drought on economic and social development. Given this, this paper examines the impact of drought shocks on firm performance by bringing the agricultural sector and the firm sector into the same analytical framework, to obtain a comprehensive picture of the impact of drought shocks on firm performance. Specifically, the study explores the impact of drought shocks on firm performance using the generalized random forest model (also known as "causal forest"). This paper focuses mainly on micro, small, and medium-sized enterprises (MSMEs), as they are one of the groups vulnerable to climate shocks. Further, the study verifies the mechanisms of market demand and input cost effects on firm performance in terms of farmers' demand and firms' input costs. Subsequently, the study examines the effectiveness of different strategies in coping with the impact of drought shocks on firms.

The contribution of this paper can be reflected in the following aspects: first, compared with the existing studies that examine the impacts of climate shocks on agriculture or firms alone, this paper incorporates farmers and firms into the same analytical framework, explores the effects of drought shocks on firms' performance, and examines the possible transmission mechanisms, which not only enriches the existing studies but also provides new perspectives for researchers to further understand the impacts of climate change on firms' performance. Second, this paper verifies the causal effect of drought shocks on enterprises from the perspectives of average treatment effect and heterogeneity treatment effect based on the generalized random forests (GRFs). Compared with the traditional econometric model, which is affected by unavoidable factors such as the model setup form, and omitted variables, the results obtained by the machine learning model are less affected by the above limitations, and thus more robust. Besides, the GRF can not only obtain the specific values of the causal effects but also calculate the robust standard errors and confidence intervals, which makes the results more interpretable. Third, we have explored the effectiveness of different adaptive strategies to cope with drought shocks based on heterogeneous causal effects, which provides a reference for actively responding to natural disasters caused by climate change and contributes to the achievement of the sustainable development goals.

2. Literature review

First, the literature related to this paper is the direct impact of natural disasters on economics development, which mainly includes research on the impact of natural disasters on farm households. Abundant studies confirm that natural disasters have a direct impact on agricultural production, affecting the income of farm households and thus negatively affecting the level of household consumption, inequality, human capital, well-being, and poverty (Paudel and Ryu 2018; Chavas et al. 2022; Cui and Zhong 2024; Liu and Wu 2024; David et al. 2021).

However, the findings of studies on the impact of natural disasters on firms are inconclusive. The results of existing studies are categorized into two types, one of which is that natural disasters hurt enterprise performance while the other finds that natural disasters promote enterprise performance. Researchers who support the negative impact of natural disasters on firm performance argue that natural disasters can negatively affect firm performance either directly through physical capital, input costs, and other channels or indirectly by affecting the industrial chain (Arrighi et al. 2022). For example, Clò et al. (2024) examined the impact of landslides and floods on firm performance based on propensity score matching–difference in differences (PSM-DID) and found that increasingly frequent hydrogeological disasters damaged firms' buildings and stock of goods and exacerbated the probability of local firms' exit from the market; with surviving firms as the subject of the study, firms' revenues and employment size both declined significantly in the 3 years following the disaster. This negative effect is more pronounced for firms with higher investments in physical assets, firms located in cities with lower-quality government, and firms controlled by nongovernmental organizations (Pan and Qiu 2022). Similarly, Zhou and Botzen (2021) explored the impact of storms and floods on firms' performance using generalized moment estimation (GMM) and found that the natural disasters have a significant negative impact on firms' sales. Based on the establishment of the negative impact of natural disasters on enterprise performance, the study concluded that different enterprises have heterogeneous effects on natural disasters, and the magnitude of the effect depends on factors that are closely related to the resilience of enterprise and regional resilience (Boudreaux et al. 2023).

In contrast to the opinion that natural disasters have a negative impact on firm performance, some scholars, based on the theory of destructive creation, argue that natural disasters lead to capital stock replacement and upgrading, stimulate innovation, and enhance firm productivity, which in turn promotes firm performance (Noy and Vu 2010; Kirchberger 2017). Coelli and Manasse (2014) utilize flood shocks as a quasi-natural experiment to estimate the impact of floods on the short-term performance of manufacturing firms based on DID. The results show that floods positively affect the growth of firms' product value added over 2 years of flooding and that firms that receive financial assistance grow faster. However, the positive effect of disasters is not always consistent, and the magnitude of its impact changes with firm size and firm productivity; thus, the positive impact of natural disasters on firms has a heterogeneous effect (Basker and Miranda 2018). Further, an investigation has emphasized that climate disasters have heterogeneous growth effects on firm sales, employment size, and capital efficiency across industries (Okubo and Strobl 2021). Moreover, some studies have found that natural disasters are expected to increase human capital to replace lost physical capital, thereby contributing to economic growth and ultimately increasing investment in physical capital and firm output growth (Skidmore and Toya 2002). Therefore, natural disasters may trigger a reallocation of labor across industries in the short run, leading to a significant and persistent wage premium between wages in different industries (Kirchberger 2017).

Other literature related to this paper are the empirical studies of heterogeneous causal effects based on GRF. Due to the ability of GRF to include more flexible functional forms, weaker premise assumptions, and the ability to estimate robust heterogeneity treatment effects, their applications in fields such as economics, sociology, and medicine have become increasingly widespread (Athey et al. 2019). Existing literature confirms either through simulated data or through empirical tests that the use of causal random forests in machine learning for causal identification and heterogeneity analysis of policies yields more accurate causal effects compared to traditional causal effects models such as difference-in-differences (DID), PSM-DID, randomized controlled trial (RCT), and regression discontinuity design (RDD) (Athey and Wager 2019).

Above all, existing studies have investigated the impacts of natural disasters on the firm performance based on econometric methods. However, most of the existing studies have examined the effects of natural disasters such as floods, earthquakes, and landslides on firm performance, while few studies have examined the effects of drought on firms in the context of the increasing frequency and intensity of droughts. Besides, few studies have explored how to effectively address the adverse impacts of climate shocks on firms in the context of the increasing prominence of climate change. Further, most of the studies considered only the average treatment effect of natural hazards without examining possible heterogeneous effects, which may have obscured the original interesting results. When examining the impact of natural disasters on firms, existing studies isolate the farm sector, and few studies have comprehensively examined how natural disasters affect both the farm sector and the firm sector. Additionally, existing studies use econometric models with biased settings or endogeneity of the selected variables, resulting in biased results. Given this, this paper integrates the farmer and firm sectors into a unified analytical framework, examines the impacts and mechanisms of drought shocks on firm performance based on GRF models from the perspectives of average and heterogeneous treatment effects, and explores possible coping strategies to mitigate the impacts of drought shocks, with a view to formulating policies for the government and the business sector to cope with climate shocks.

3. Theoretical frameworks

a. *Drought shocks and firm performance*

In this paper, farm households and firms are incorporated into the same analytical framework. Specifically, drawing on existing studies (Grabrucker and Grimm 2021), a classical household model is used to analyze and incorporate drought shocks into the production technology function. It is assumed that the corporate sector makes production decisions based on profit maximization, prices are exogenous, and farm households engage in agricultural production as well as nonfarm employment to generate income. The price of agricultural products is p_f and the output is Q_f; the price of agricultural production factor costs is p_{K_1} and the quantity of inputs is K_1; the price of farmers' labor is w and the supply of their agricultural production and nonfarm labor is L_f、L_n respectively. For farmers, there are

$$\max E\pi_f = p_f Q_f - p_{K_1} \times K_1 - wL_f + wL_n. \tag{1}$$

For agricultural production Q_f, which is assumed to be subject to a Cobb–Douglas (C-D) functional form, it is jointly affected by the level of agricultural technology and drought shocks. Specifically, there is an agricultural production technology function:

$$Q_f = A\theta K_1^{\alpha} L_f^{\beta}, \tag{2}$$

where A is the technology level, θ is the drought shock, and α and β are the constants. Based on Eqs. (1) and (2), we can obtain the parsimonious production function $Q_f = Q_f(p_f, p_{K_1}, w; A; \theta)_o$. Further, farm households maximize their utility for agricultural products, firm sector products, and leisure consumption under income constraints. Specifically,

$$\mathrm{Max}EU(C_f, C_e, C_l) = C_f^{\gamma_1} \times C_e^{\gamma_2} \times C_l^{\gamma_3};\ p_f C_f + p_e C_e + wC_l \le E\pi_f, \tag{3}$$

where C_f, C_e, and C_l are the number of agricultural products consumed by farmers, the number of products produced by the enterprise, and the leisure time respectively; γ_i is the parameter of the utility function; and $\gamma_1 + \gamma_2 + \gamma_3 = 1$, $\gamma_i > 0$. The utility function is concave and monotonically nondecreasing. Based on Eq. (3) and the Lagrange multiplier method, we can get the consumption function $C_e = \gamma_2 E\pi_f/p_e$ under the goal of maximizing the utility of farmers to the enterprise, and $\partial C_i/\partial E\pi_f = \gamma_i/p_i > 0$. Combined with Eq. (1), we can get the parsimonious function of the demand of farmers to the enterprise's products $C_e = C_e(p_f, p_e, w, E\pi_f)$.

For the firm sector, the price of the firm's product is p_e, and the quantity demanded is Q_e; the price of the firm's factor costs is p_{K_2}, and the quantity of inputs is K_2; and the quantity of hired labor required for the firm's production is L_e; from this, the firm's production function can be obtained as follows:

$$\max E\pi_e = R_e - C_{K_2} - C_{L_e} = p_e Q_e - p_{K_2} \times K_2 - wL_e. \quad (4)$$

Similarly, assuming that the production technology of the firm's sector follows the C-D functional form, we can get the expression Q_e as follows:

$$Q_e = BK_2^{\alpha} L_e^{\beta}, \quad (5)$$

where B is the firm's productivity level. Based on Eqs. (4) and (5), the firm output parsimony function can be obtained as $Q_e = Q(p_e, p_{K_2}, w; B)$, with $Q_e = C_e + C_u = \delta_1 Q_e + \delta_2 Q_e$, $\delta_1 + \delta_2 = 1$, where C_u is the consumption of the firm's products by the urban residents. Similarly, the enterprise employment function can be obtained as $L_e = (Q_e/B)^{1/(\alpha+\beta)} \times (p_{K_2}\beta/w\alpha)^{\alpha/(\alpha+\beta)}$, whose parsimony is $L_e = L_e(Q_e, p_{K_2}, w; B)$; and the factor function $K_2 = (Q_e/B)^{1/(\alpha+\beta)} \times (w\alpha/p_{K_2}\beta)^{\beta/(\alpha+\beta)}$.

Next, we derive the hypotheses of this paper based on the above model in terms of farmers' demand for firms' products and firms' input costs.

Existing research results confirm that drought shocks significantly reduce agricultural productivity, increase the incidence of crop pests and diseases and thus increase the cost of agricultural inputs, reduce agricultural production, and ultimately will significantly reduce the income of farmers (IPCC 2022; Liu and Wu 2024). Therefore, $\partial Q_f/\partial\theta < 0$, $\partial E\pi_f/\partial\theta = (\partial E\pi_f/\partial Q_f) \times (\partial Q_f/\partial\theta) < 0$. Since the income of the farmers decreases after the drought shock, this will significantly affect the consumption of the firm's products by the farmers, i.e., there is $\partial C_e/\partial\theta = (\partial C_e/\partial E\pi_f) \times (\partial E\pi_f/\partial\theta) < 0$. Therefore, the local farmers' demand for the products produced by the firms decreases when the region is exposed to the drought shock, i.e., $C'_e < C_e$. According to the theory of price stickiness, it is difficult for enterprises to adjust the price of their products in a short time when they are hit by a drought; thus, the price of their products p_e remains unchanged, and the supply function of the enterprise remains unchanged. Since urban residents mainly rely on nonfarm employment income, its income is not affected by the drought shock, so in the case of enterprise, product prices remain unchanged C_u. Therefore, we can get the total demand for enterprise products under drought shock as $Q'_e = C'_e + C_u < Q_e$, i.e., the total demand of residents for enterprise products is reduced under drought shock, and $\partial Q_e/\partial\theta < 0$.

According to the demand-supply theory, if the supply function is unchanged, the demand decreases, the sales of the enterprise is expected to decrease, and the revenue of the enterprise decreases, i.e., $\partial R_e/\partial\theta = \partial p_e Q_e/\partial\theta = p_e(\partial Q_e/\partial\theta) < 0$. For the employees hired by the firm, when the firm is exposed to drought shock, there are $\partial L_e/\partial\theta = (\partial L_e/\partial R_e) \times (\partial R_e/\partial\theta) = (\partial L_e/\partial Q_e) \times (\partial Q_e/\partial R_e) \times (\partial R_e/\partial\theta) < 0$,[1] i.e., when facing the drought shock, the scale of hiring of the firm will be reduced. For corporate profitability, there is $\partial\pi_e/\partial\theta = p_e(\partial Q_e/\partial\theta) - p_{K_2}(\partial K_2/\partial\theta) - w(\partial L_e/\partial\theta)$; the positivity and negativity of $\partial\pi_e/\partial\theta$ are affected by the magnitude of the decrease in demand and the magnitude of the decrease in cost. Therefore, the impact of drought shocks on the profits of regional firms needs to be rigorously empirically tested. Based on the above discussion, this paper proposes the following hypothesis:

Hypothesis 1: Drought shocks have a negative impact on firms' revenues and employee employment scale, and their impact on firms' profits is uncertain.

Hypothesis 2: Drought shocks affect firms' revenues and employment mainly through the demand effect, which, together with the cost effect, affects firms' profitability.

b. *Adaptive strategies for drought shocks*

To effectively respond to drought shocks, the government's efforts to improve agricultural infrastructure can effectively mitigate the negative impacts of drought shocks on farmers' incomes and agricultural production (Kühl Teles and Cesar Mussolini 2012). Existing studies have shown that, on the one hand, improvements in agricultural infrastructure can enhance the efficiency of agricultural production, which boosts farm household income (Hamilton et al. 2022). On the other hand, agricultural infrastructure improvement can also promote the development of the local nonfarm economy and enrich the source of income of farm households. Therefore, a strong promotion of infrastructure not only mitigates the negative impacts of drought shocks on agricultural yields but also stabilizes farm household incomes. Besides, agricultural technology upgrading is also seen as an effective measure to mitigate the adverse effects of climate change (Conley and Christopher 2001). Taking agricultural mechanization as an example, the adoption of agricultural mechanization by farmers can effectively release human resources for other nonfarm activities and enhance household income (Afridi et al. 2023). In the face of climate shocks, the use of mechanization can quickly carry out agricultural replanting or harvesting activities, thus minimizing farmers' losses.

Promoting regional financial inclusion is also an effective strategy for coping with drought shocks. Firms need finance to invest in technology to build resilience to shocks (Gannon et al. 2022). However, for most MSMEs, access to formal finance is limited due to inherent challenges such as information asymmetry, high-risk premiums, and monitoring costs (Quartey et al. 2017). Constraints on access to formal finance force firms to adopt unsustainable coping mechanisms,

[1] $\dfrac{\partial L_e}{\partial Q_e} = \dfrac{1}{B}(Q_e/B)^{(1-\alpha-\beta)/(\alpha+\beta)} \times \left(\dfrac{p_{K_2}\beta}{w\alpha}\right)^{\alpha/(\alpha+\beta)} > 0,\ \dfrac{\partial Q_e}{\partial R_e} = \dfrac{1}{p_e} > 0.$

such as downsizing production, thereby limiting growth opportunities (Shibia 2024). In contrast, inclusive finance effectively mitigates the information asymmetry between firms and financial intermediaries, simplifies the lending process, and effectively promotes the growth of firms, especially small and micro enterprises. Similarly, farmers can also use digital inclusive finance for credit, entrepreneurship, and part-time jobs, which effectively stabilizes their household income and consumption levels (Liu and Guo 2023), thus ensuring that overall consumption demand is less affected by drought shocks. Based on this, this paper proposes the following hypothesis;

Hypothesis 3: We can respond to the impact of drought shocks on firms by promoting the implementation of agricultural infrastructure, upgrading the level of regional agricultural mechanization, and enhancing the level of digital financial inclusion.

4. Research design

a. Empirical model

Previous research has utilized DID mostly to estimate the causal effect of natural disasters on firms' performance. However, the accurate use of the aforementioned method requires the fulfillment of stringent preconditions, namely, the parallel trend assumption and the no confounding assumption. This assumption is usually difficult to establish, first because of the selectivity of the treatment and second because of the variability of the sample. To find samples that satisfy the criteria applicable to DID, researchers often use propensity score matching methods to obtain samples that fit the analysis. However, during the matching process, self-matching and loss of sample size are issues that cause the estimates to remain biased. Besides, utilizing the DID to obtain the average treatment effect of the samples cannot obtain the heterogeneity treatment effect. However, section 3 points out that the impact of drought shocks on firm performance will vary from firm to firm. Although this has been recognized in many previous related studies, all existing studies have estimated only their average treatment effects based on traditional econometric methods, and few studies have examined heterogeneous treatment effects. In contrast, our approach can be based on the conditional average treatment effect (CATE), thus obtaining individualized causal effects.

According to the Rubin causal model (Rubin 1974), there are two potential outcomes, Y^0 and Y^1. Assume that there are observations (X_i, Y_i, D_i), where $i = 1, \ldots, n$, and n is the number of observations, $X_i = x \in R^p$ is the p-dimensional feature vector, $Y_i \in R$ is an outcome variable, and $D_i \in \{0, 1\}$ is a treatment variable, i.e., whether or not the firm is subject to a drought shock. For each firm with feature X, there is a conditional average treatment effect: $\tau(x) = E(Y_i^1 - Y_i^0 | X_i = x)$, where Y_i^1 refers to the outcome variable that is subject to the shock. However, following Holland (1986), we cannot observe both Y_i^1, Y_i^0. Based on Chernozhukov et al. (2018), Athey et al. (2019) argue that the CATE can be estimated by a simple model: $Y_i = \tau(x)D_i + m(x) + \varepsilon_i$. Transforming this to residual–residual regression (Chernozhukov et al. 2018), we obtain the estimate $Y_i - \hat{m}(x) = \tau(x)[D_i - \hat{e}(x)] + \varepsilon_i$, where $\hat{e}(x) = P(D_i = 1| = X_i = x)$ is the propensity score, $\hat{m}(x) = E(Y_i = y|X_i = x)$ is the expected outcome, and ε_i is the residual term.

Therefore, to obtain CATE$[\tau(x)]$, $m(x)$, $e(x)$ must be estimated. Athey et al. (2019) developed the generalized random forest model to obtain the parameters of interest. The treatment effect estimate $\hat{\tau}(x)$ is obtained as

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{n} \alpha_i(x)(\widetilde{D_i} - \overline{D_\alpha})(\widetilde{Y_i} - \overline{Y_\alpha})}{\sum_{i=1}^{n} \alpha_i(x)(\widetilde{D_i} - \overline{D_\alpha})^2}. \tag{6}$$

Among them, $\alpha_i(x) = (1/B)\sum_{b=1}^{B} I\{[X \in L_b(x), i \in S_b]\}/|L_b(x), i \in S_b|$, where B is the amount of trees, $L_b(x)$ is the leaf of the bth tree containing the test point x, and S_b is the subsample used to grow the bth tree; $\widetilde{D_i} = D_i - \hat{e}(x)^{\text{oob}}$, $\widetilde{Y_i} = Y_i - \hat{m}(x)^{\text{oob}}$, where $\hat{e}(x)^{\text{oob}}$ and $\hat{m}(x)^{\text{oob}}$ are the propensity scores and expected outcomes obtained by out-of-bag prediction[2]; $\overline{D_\alpha} = \sum_{i=1}^{n} \alpha_i(x)\widetilde{D_i}$, $\overline{Y_\alpha} = \sum_{i=1}^{n} \alpha_i(x)\widetilde{Y_i}$. It can not only obtain the heterogeneity treatment effect but also effectively alleviate the unavoidable difficulties such as confounding factors and omitted variables in the econometric statistical methods, so its results are more robust and reliable. We will verify this in the robustness test.

b. Omitted variable and confounding factors

It is without doubt a primary drawback of existing identification strategies that causal effects can only be identified if the researcher observes all relevant confounders (see also Fig. S1 in the online supplemental material). Otherwise, the estimates are biased because unobserved omitted variables are associated with treatments and outcomes (Wager and Athey, 2018). This seems to be an unavoidable problem. However, GRF can address endogeneity bias caused by unobserved confounders, although we do not directly include all potential confounders (Wang and Blei 2019). Specifically, as shown in Fig. 1, if the model does not control for all observed confounders from X_1 to X_K and unobservable confounders U_1, U_2, the effect of outcome Y cannot be correctly identified. Since U_1 and U_2 are unobservable, it is not possible to directly control these confounders. However, assuming that observed and unobserved confounders are correlated ($U_{1,2} \rightarrow X$) and that unobserved confounders are reflected in a complex, nonlinear, and high-dimensional combination of a large number of observed confounders (the potential confounder space), it is possible to capture variation from the unobserved confounder space, if the causal forest accurately maps the potential confounder space. In this case, we can close the backdoor path $W \leftarrow U_{1,2} \rightarrow Y$ and identify the causal effects of drought shocks on firm

[2] Note: OOB denotes out-of-bag predictions, i.e., these predictions are generated by using only the portion of trees that do not have that data point in the respective subsample used to generate the predictions.

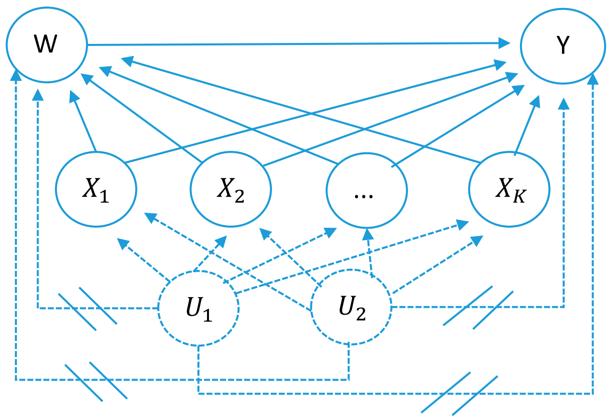

FIG. 1. Directed acyclic graph with unobservable factors.

performance.[3] Therefore, high-dimensional nonlinear, highly complex combinations of observed features X can serve as an approximation of unobserved confounders and can represent the potential covariate space to some extent, as long as all relevant information is hidden in the observed data. Such a characterization is not possible with traditional regression techniques, and thus cannot adequately approximate the latent space. However, the GRF can capture variation from this unobserved confounder space through their complex structure (Bennett and Kallus 2019). Thus, the ability of GRF to approximate sufficiently omitted variables is seen as a way to address unbounded conditions.

c. Data

The precipitation data used in the research are obtained from the China Meteorological Data Network, which is the most representative and reliable provider of meteorological data in China by providing annual, monthly, and daily data for all cities in China. The study obtains monthly precipitation shocks for each prefecture-level city in the system and then matches enterprise data and household data with the city code. The enterprise data and household data used in this paper come from the China Household Finance Survey (CHFS) database, which covers the comprehensive financial and economic status of households in 29 provinces of China and is broadly representative. The sampling scheme of the CHFS utilizes a stratified, three-stage proportional to size measure (PPS) sampling design. We selected data from the questionnaires in the production and operations section of the CHFS database, which surveys the production and operations of regional firms. We removed samples containing missing values. Since the firms' questionnaires were added to the CHFS database starting in 2015, the database is currently updated to 2019. Thus, we finally obtain a total of 11 531 firm and 14 410 farm annual observations for the three periods of 2015, 2017, and 2019, where firms' industries mainly include agriculture, retail trade, construction, services, transportation, leasing, and other industries. The above sample covers 211 cities in China. Besides, the regional characteristic variables were obtained from the China Statistical Yearbook.

d. Variables

1) Drought shock

Based on existing research practices, we use the standardized precipitation index (SPI) proposed by the International Meteorological Organization to measure drought shocks. The SPI is calculated as follows.

First, a gamma distribution is fitted based on the precipitation data with density function $g(x)$; then, the cumulative probability $(x) = \int_0^x g(t)dt$, where x is the precipitation amount. If $x = 0$, $G(x = 0) = m/n$, where m is the number of samples with zero precipitation and n is the total sample size. Next, $G(x)$ is transformed to SPI using the inverse standard normal distribution with

$$\mathrm{SPI} = S\left\{t - \frac{(c_2t + c_1)t + c_0}{[(d_3t + d_2)t + d_1]t + 1}\right\}, \tag{7}$$

where $t = \sqrt{\ln[1/G(x)^2]}$; if $G(x) > 0.5$, $S = 1$; otherwise, $S = -1$; $c_0 = 2.515517$, $c_1 = 0.802853$, $c_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, and $d_3 = 0.001308$. We use R software for calculations to obtain the SPI, and then based on China's national meteorological drought standardized, we define the drought shock as $ds = 1$ if SPI < -0.5. Consistent with existing studies (Khalili et al. 2021), this paper investigated the impacts of drought shocks in the growing season (April–September) on the firm performance, mainly from prefecture level. Therefore, the result is estimated using SPI values at the city level on a 6-month scale.

2) Outcome variables

This paper focuses on the impact of drought shocks on firms' performance, and based on the theoretical hypotheses in Part III, the outcome variables we selected are firm sales, employment, and profit.

3) Characterization variables

To control potential confounders and obtain more precise causal effects, referring to existing studies (Clò et al. 2024), we control for both firm and regional characteristics. Among them, firm characteristic variables include firm age, firm assets, firm asset-liability ratio, firm organization form, firm nature, and the industry in which the firm is located; area characteristic variables include economic growth, regional population size, financial development level, general budget income, and human resource level. In the mechanism analysis, the mechanism variables mainly include the consumption level of farm households and enterprise input costs. In testing the mechanism of market demand effect, the study also controls for household characteristics, mainly including household size, assets, age of the head of household, marital status, physical condition, and education level. In the adaptive strategies analysis, the adaptive strategies variables mainly include effective irrigated area, total

[3] For the purpose of brevity, links between confounding factors are not shown.

TABLE 1. Variables definitions and descriptive statistics.

	Variables	Definitions	Mean	Standard deviation
Outcome variables	Sales	Enterprise revenue/10,000 (million yuan)	28.68	93.81
	Employment	Total number of employees of the enterprise	4.082	14.82
	Profit	Corporate profit/10,000 (million yuan)	6.853	31.99
Treatments variable	Drought shocks	Precipitation below the long-term average	0.1933	0.4939
Firm characteristics	Firm e-code	Firm identifier		
	Age of the firm	Length of firm establishment (years)	10.13	9.018
	Firm organization form	Joint stock limited company = 1, limited liability company = 2, partnership = 3, sole proprietorship = 4, self-employed/ business owner = 5, and no formal form of organization = 6	3.832	1.814
	Industry	Industry in which the firm is located	3.900	2.261
	Assets	Total firm assets/10,000 (yuan)	51.82	379.4
	Assets and debt ratio	Enterprise debt/enterprise assets	0.3282	0.2813
Regional characteristics	Economic development	GDP growth rate	7.561	2.6568
	Economic scales	GDP per capita (million yuan)	6.868	3.7382
	Population size	Resident population (10 000 people)	722.2	606.3
	Financial development	Loan balance of regional financial institutions/GDP	1.8050	0.7778
	Industrial structure	Tertiary industry GDP ratio	48.06	10.31
	Public service	Ratio of local government fiscal expenditure to GDP	0.1870	0.0869
	City code	City identifier		
	Year	Dummy variable		
Mechanism variables	Total income	Total household income/10,000 (million yuan)	4.343	9.717
	Total consumption	Total household consumption/10,000 (10,000 yuan)	3.629	4.525
	Subsistence consumption	Total food and daily household consumption/1000 (1000 yuan)	20.52	26.10
	Developmental consumption	Healthcare and education consumption/ 1000 (1000 yuan)	8.337	19.91
	Recreation consumption	Tourism and entertainment consumption/ 1000 (1000 yuan)	5.933	20.87
	Profitability	Enterprise profit/enterprise revenue	0.3411	0.2917
	Per capita consumption in the region	Per capita consumption of rural inhabitants in the region (10,000 dollars)	1.1953	6.7540
Adaptive strategy variables	Effective area	Effective irrigated area of the region/total cultivated area	0.3799	0.1504
	Mechanization	Total regional mechanized power/cultivated land area	0.5072	0.3205
	Digital finance	Regional digital financial inclusion index/100	2.052	4.0700

mechanization momentum per unit of cultivated area, and digital financial inclusion index. Furthermore, we include year and enterprise dummy variables in all models. The specific definitions and descriptive statistics of each variable are shown in Table 1, where the descriptive statistics of household characteristics variables are shown in Table S1.

5. Results

a. *Baseline results*

Based on the GRF model, the results of the impact of drought shocks on firm performance are obtained by incorporating firm characteristics and regional characteristics as shown in Table 2. Columns (1)–(4) in Table 2 show the magnitude of the average causal effect of drought shocks on firms' sales without clustering and with clustering at different levels, respectively. We can find that regardless of the level at which the model is clustered, the results show that drought shocks have a negative impact on firms' revenues. Further, to select the model that measures the causal effect optimally, we calibrate the causal model with the test. Specifically, we compute the mean forest prediction (MFP) statistic for each model; the closer its value is to 1 and significant at the statistical level, the better the model quality (Chernozhukov et al. 2025). It can

TABLE 2. Benchmark results. Note: Standard errors are in parentheses. The triple asterisk (***) indicates significance at the 1% level, double asterisk (**) indicates significance at the 5% level, and asterisk (*) indicates significance at the 10% level.

	Sales				Employment				Profit			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Drought shocks	−8.299***	−6.684***	−13.30***	−7.279***	−1.4829***	−1.2857***	−0.8454***	−0.7539***	−3.691***	−1.7668***	−1.664	−1.7165***
	(2.256)	(2.1004)	(1.418)	(1.681)	(0.2688)	(0.6603)	(0.3984)	(0.2857)	(0.8483)	(0.8511)	(0.5158)	(0.4916)
Firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustering	No	Firm	Industry	City	No	Firm	Industry	City	No	Firm	Industry	City
N	11 531	11 531	11 531	11 531	11 531	11 531	11 531	11 531	11 531	11 531	11 531	11 531
Model	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest	Causal forest
MFP	2.069	1.926	0.7708**	1.232***	0.4439	1.6224***	1.386	1.089**	0.7358(	1.739	1.202	1.166***
	(1.318)	(1.345)	(0.4971)	(0.3129)	(0.7546)	(0.1314)	(0.9943)	(0.069)	1.188)	(1.391)	(1.463)	(0.4523)
DFP				1.397***(0.370)				0.9196***				1.329***
								(0.2857)				(0.4206)

be noticed that the causal random forest model clustered at the city level has the best quality. The following sections are analyzed based on the optimal model setting.

Columns (5)–(8) and (9)–(12) of Table 2 report the causal effects of drought shocks on firm employment and firm profits without clustering as well as with clustering at the different level, and the results show that drought shocks have a negative effect on firm employment and firm profits. Similarly, the MFP statistics of the different models are examined and the model clustered at the city level is found to be optimal. Based on the analysis of the results of the optimal model, it is found that compared with the enterprises that are not affected by the drought shock, being affected by the drought shock will lead to an average reduction of 72,000 yuan in revenue, 0.75 employees, and 17,000 yuan in profits, and that hypothesis 1 of this paper is valid. On top of calculating the average treatment effect, we calculated the differential forest prediction (DFP) statistic of the optimal model, whose value, if significantly greater than 0, indicates that the drought shock has a significant heterogeneous effect on enterprise performance. The results of the DFP statistic indicate that drought shocks have a heterogeneous impact effect on firm performance.

b. *Robustness checks*

1) RESULTS VALIDITY CHECKS

The validity of the causal inference results depends on whether the sample satisfies the common support assumption and the Stable Unit Treatment Value Assumption (SUTVA) assumption. To test whether the results of this paper satisfy the common support assumption, we followed Liu and Guo (2023) and used the nearest neighbor propensity score matching to match the samples that suffered from drought shocks (treatment group) with those that did not (control group). The results of the balance test after matching are shown in Table S2, which indicates that there is no significant difference between the matched samples and the matching results are satisfactory. Estimation based on the new samples after matching is performed using GRF and the results obtained are shown in Table 3. The results in columns (1)–(3) of Table 3 indicate that the results obtained based on the matched samples are not significantly different from the optimal benchmark model. Further, Figs. S2 and S3 demonstrate the distribution of propensity scores before and after matching. The results show that the propensity scores of both shocked and unshocked firms before and after matching are in the range of 0–1, which satisfies the common support assumption, which indicates that the results of the benchmark model are robust and reliable as well as the results after matching.

Subsequently, we check whether the results of this paper are affected by the SUTVA assumption, which requires that there is no correlation between the samples in the treatment group. Considering the correlation among the shocked firms, the optimal benchmark model in this paper clusters the samples to the city level when estimating the results, which effectively ensures the reliability of the results. To further isolate the interference of correlation among shocked firms in the results of this paper, all firms in each city are aggregated to the

TABLE 3. Results validity checks. Note: Standard errors are in parentheses. The triple asterisk (***) indicates significance at the 1% level.

Variables	PSM			Sample merged to city level		
	(1) Sales	(2) Employment	(3) Profit	(4) Sales	(5) Employment	(6) Profit
Drought shocks	−7.3604***(1.6533)	−0.7602***(0.2488)	−1.7252***(0.5275)	−7.3781***(1.7651)	−0.7622***(0.2790)	−1.7301***(0.5316)
Firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes
N	8530	8530	8530	663	663	663

city level, i.e., there is only one firm in each city, similar to the practice of federal causal inference. Specifically, for city i's firm indicator j, there is $\widetilde{X}_{ij} = (n_i/N)\overline{X}_{ij}$, where $\widetilde{X}_{ij}$ is the indicator j aggregated to the level of city i, n_i is the number of original firms in city i, N is the amount of all the firms in the sample, and $\overline{X}_{ij}$ is the mean value of the indicator j for all firms in city i. Following the above processing, there is only one firm in each city after the merger, in which case there is no correlation between the shocked firms. Moreover, the mean value of each indicator of the merged enterprises is equal to the mean value of each indicator of the original sample size of N, which is easy to compare the estimation results. Estimation based on the merged sample using GRF obtains the results shown in Table 3. The results show that the results obtained based on the merged sample are consistent with the results of the benchmark regression and the results are not affected by the SUTVA assumption and are robust.

Further, we draw a random sample from each city affected by the drought shock and then use the drawn sample for estimation, with the above operation resulting in no correlation between treatment groups. We repeat the above operation 1000 times to ensure that each shocked firm has a chance to be sampled. The results obtained based on the above steps are shown in Fig. S4. The results are not significantly different from the benchmark results and the paper is not subject to the SUTVA assumption and the results are reliable. The above results can demonstrate that the premise assumption of uncorrelation between treatment groups is greatly weakened due to the fact that the GRF generates similar samples to a single treatment group from the control group, thus obtaining a single treatment effect. Thus, the GRF can obtain precise results and the superiority of using the GRF for causal inference is highlighted.

2) OMITTED VARIABLES AND OTHER CONFOUNDING FACTORS

To effectively mitigate omitted variable bias, we rely on the assumption that all relevant information is latent in the observed data. To test the sensitivity of this potential covariate assumption, we assume, based on a counterfactual approach, that our data do not contain enough information and that the existing feature space struggles to capture the characteristics of the omitted variables. Therefore, at this point, adding additional features to the model or reducing the feature variables in the original feature space would drastically change the original results.

Specifically, we conducted the following tests. First, we add other variables that may affect the performance and we include the way of running the business in the feature space because the way of running the business may have an impact on the sales and hiring of the firm. Further, we continue to add the feature of the amount of tax paid by the firm, which measures the level of revenue and profit. For how to remove the feature space, we refer to existing studies and calculate the degree of importance of each variable to firm performance. The specific results are shown in Fig. S5, which indicates that firm assets are the most important variable affecting business performance. For this reason, we remove this feature from the test. Based on the above steps, samples containing

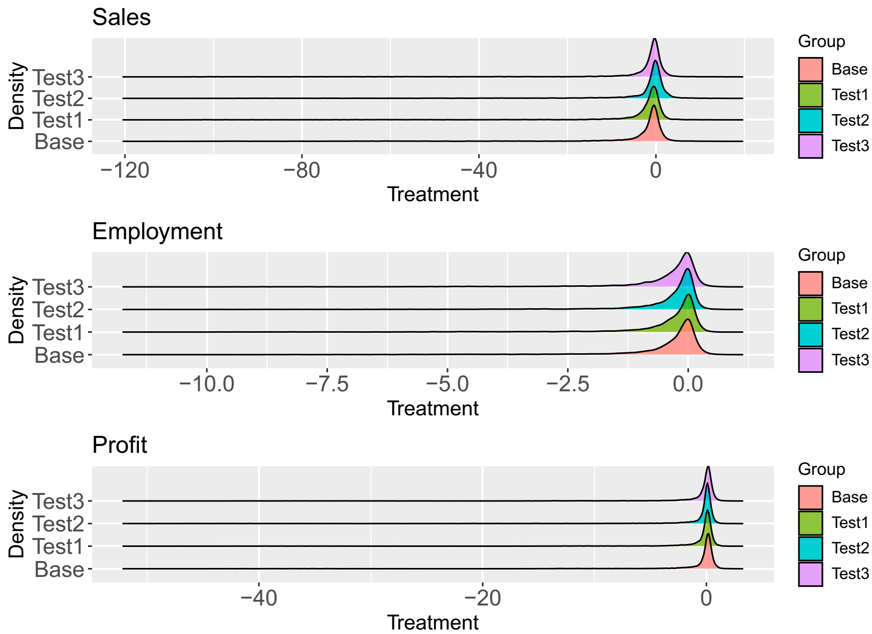

FIG. 2. Data informativeness test.

different feature variables are obtained to retrain the model and the results are obtained as shown in Fig. 2. From the results in Fig. 2, it can be found that the results reobtained based on the new feature space are not significantly changed compared to the baseline results, neither by adding confounders nor by deleting important variables. This proves that the research data contain enough information that affects firm performance, the results are unlikely to be affected by omitted variables, and the model used in the study can effectively overcome the bias of the results caused by unobserved confounders and omitted variables, so the advantages of the data-driven approach can be realized.

3) ADDITIONAL ROBUSTNESS CHECKS

To further determine the robustness of the results, we also performed the following robustness checks: (i) changing the calculating method of drought shock; (ii) consider sharp transition between drought and flood shocks; (iii) consider regional differences; (iv) repeated shock events; (v) spillover effects; (vi) nonrandom distribution of firms. Due to word limitations, the exact results are presented in the supplemental material. The results show that the robustness of our results holds.

6. Expanded analysis

a. *Mechanism analysis*

1) MARKET DEMAND EFFECTS

The theoretical analysis in Part III shows that drought shocks adversely affect household consumption demand by lowering farm household incomes, and that a decline in household consumption demand in the region leads to a downturn in market demand, which in turn adversely affects enterprise revenues and employment size. Therefore, we examine the impact of drought shocks on farm household income and consumption. First, we examine the impact of drought shocks on households' total incomes in 2015, 2017, and 2019. In the model training process, we add household characteristics and head of household characteristics to the baseline model setup and obtain the results shown in column (1) of Table 4. It can be found that drought shocks will significantly reduce the income of regional farmers. Following the examination of household incomes, we further explore the impact of drought shocks on the consumption structure of households.

The impact of drought shocks on farmers' consumption is shown in Table 4. The results show that drought shocks have a significant negative impact on the total consumption of the household and the different types of consumption. The average total consumption of households in drought-shocked areas will be reduced by 12,000 yuan compared with that of nonshocked households. Further, we examine the impact of drought shocks on regional consumption at the macrolevel. Specifically, we explore the impact of drought shocks on per capita consumption of regional rural residents and obtain the results shown in the table. The results in column (6) of Table 4 indicate that rural consumption levels in regions hit by drought shocks are significantly lower. The decline in overall consumption by regional farmers will lead to a decrease in demand for products in the market, which will adversely affect the revenue and employment status of local enterprises.

2) INPUT COST EFFECTS

The empirical results also show that drought shocks have a negative impact on firms' profits. Based on the theoretical analysis in the third part, it can be inferred that drought shocks lead to a decrease in the demand for firms' products

TABLE 4. Mechanism analysis. Note: Total consumption is in tens of thousands of dollars and the remaining subconsumption variables are in thousands of dollars. Standard errors are in parentheses. The triple asterisk (***) indicates significance at the 1% level.

Variables	(1) Income	(2) Total consumption	(3) Subsistence consumption	(4) Developmental consumption	(5) Recreational consumption	(6) Per capita consumption in the region	(7) Profitability
Drought shock	−1.065***(0.1753)	−1.1587***(0.4051)	−1.1453***(0.2150)	−2.4001***(0.1957)	−2.9403***(0.3240)	−2.7907***(0.4302)	−0.0217***(0.0036)
Firm characteristics	Yes	Yes	Yes	Yes	Yes		Yes
Regional characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes		
N	14410	14410	14410	14410	14410	303	11531

that is greater than the decrease in firms' costs, which in turn has a significant impact on firms' profits. Since the database does not provide regional consumption data and firms' costs, to test whether the mechanism is valid, we test the effect of drought shocks on regional firms' profitability and obtain the results shown in column (7) of Table 4. The results show that drought shocks have a significant negative effect on corporate profitability. According to Eq. (4), $E\pi_e/Q_e = R_e - C_{K_2} - C_{L_e} = p_e - C/Q_e$, when the corporate profitability decreases, due to the price stickiness, p_e keeps unchanged, and the magnitude of the decrease in demand becomes larger than the magnitude of the decrease in cost. The synthesis of the previous part of the results can prove that the hypothesis 2 of this paper is valid.

b. *Heterogeneity analysis*

The results of the optimal benchmark model in Table 2 show that the DFP statistic is significantly positive, indicating that there is heterogeneity in the results of this paper. Moreover, the base group in Fig. 2 indicates that the treatment effects have a left-skewed distribution and that the heterogeneity effect exists. To confirm the existence of the above heterogeneous impact effects, we obtain the conditional treatment effect of drought shocks on firms' performance under different asset sizes and ages based on the baseline optimal causal forest model, and the specific results are shown in Fig. 3, where the shaded parts are the 95% confidence intervals of the treatment effects. From Fig. 3, it can be found that drought shocks have heterogeneous impacts on firms with different asset sizes, and the larger the asset size, the smaller the negative impacts on firms, and the impacts of drought shocks on firms will tend to be insignificant after the firm's asset size reaches a certain level. A reasonable explanation for this is that the larger the asset size, the more resilient the firm is, and the more it has the strength to seek new markets to stabilize production in response to shocks to market demand, so it is less likely to be affected by drought shocks.

It is worth noting, however, that drought shocks have a "U"-shaped effect on the employment size of firms with different assets. This can be explained by the fact that firms with small assets tend to be self-employed or family-owned and are less likely to make layoffs because a reduction in hiring would mean a shutdown of production. Thus, at a certain asset level, drought shocks have a greater impact on the employment level of firms that are larger in terms of asset size. When firms are above a certain asset size, they have enough assets to cope with changes in market demand and are less affected by drought shocks, so they are less likely to make layoff decisions after the shock. For firms at different stages of establishment, drought shocks have a greater impact on start-up firms, which are significantly affected by drought shocks in terms of revenue, employment size, and profits. This is because mature firms have more experience, social networks, and tools to cope with a decline in market demand and higher input costs due to drought shocks than start-ups, which may be more affected by drought shocks due to their lack of experience in obtaining sufficient orders to carry out production activities in the face of a decline in market demand.

FIG. 3. Analysis of heterogeneous causal effects based on firm asset size and age.

Furthermore, we examined the extent to which drought shocks affect the firm performance in different industries, and the results are shown in Fig. 4 and Figs. S6 and S7. We found that the agriculture is the most affected by drought shocks, followed by the transportation industry, while the leasing and service industry is less affected by drought shocks. The above results may be explained by the fact that drought shocks have a direct negative impact on agriculture, which is also the most affected due to the increase in the price of agricultural products. Transportation is also more affected as the demand for transportation of goods and raw materials declines in all sectors due to low market demand and rising input costs. Overall, drought shocks affect firms through market demand and input cost effects and have a heterogeneous impact on firms of different asset sizes, ages, and industries.

c. *Analysis of the effectiveness of adaptive strategies*

The results in section 5 reveal that drought shocks have a negative effect on firms' performance and adversely affect local economic development as well as the employment of the residents. How to effectively mitigate the negative impacts of extreme weather events on the economy and society in the context of the increasing frequency and intensity of such events is of paramount importance. The theoretical analysis suggests that coping strategies such as vigorously developing agricultural technology, improving agricultural infrastructure, and upgrading the level of digital financial inclusion in the region can effectively mitigate the impact of climate change. To this end, we tested whether the above coping strategies can effectively mitigate the negative impacts of drought on business performance. Specifically, we refer to the existing study (Hamilton et al. 2022), which uses the effective irrigated area per unit of arable land in the region to measure the construction of agricultural infrastructure, and the total power of agricultural machinery per unit of arable land in the region to measure the agricultural mechanization of the region, and calculates the conditional mean treatment effects of the above adaptive strategy variables based on the benchmark causal forest model, and obtains the results shown in Fig. 5.

From the results in Fig. 5, with the increase in the effective irrigated area per unit of arable land in the region, the level of mechanization per unit of arable land in the region, and the level of digital financial inclusion, we found that the negative impacts of drought shocks on firms' revenues, hiring, and profits are mitigated. This suggests that the above coping strategies are effective in mitigating the negative impacts of drought shocks on firms. On the one hand, with the improvement of regional agricultural infrastructure and mechanization, farmers in the region can effectively cope with the negative impacts of climate shocks and mitigate crop losses, which effectively reduces the input costs of enterprises that use crops as a means of production and stabilizes farmers' household incomes and consumption levels. On the other hand, improving the level of digital inclusive finance in the region can help ease the credit constraints of farmers and enterprises, promote the consumption of local residents, and stabilize the production of enterprises, which can also be used as an effective measure to cope with climate change. Taken

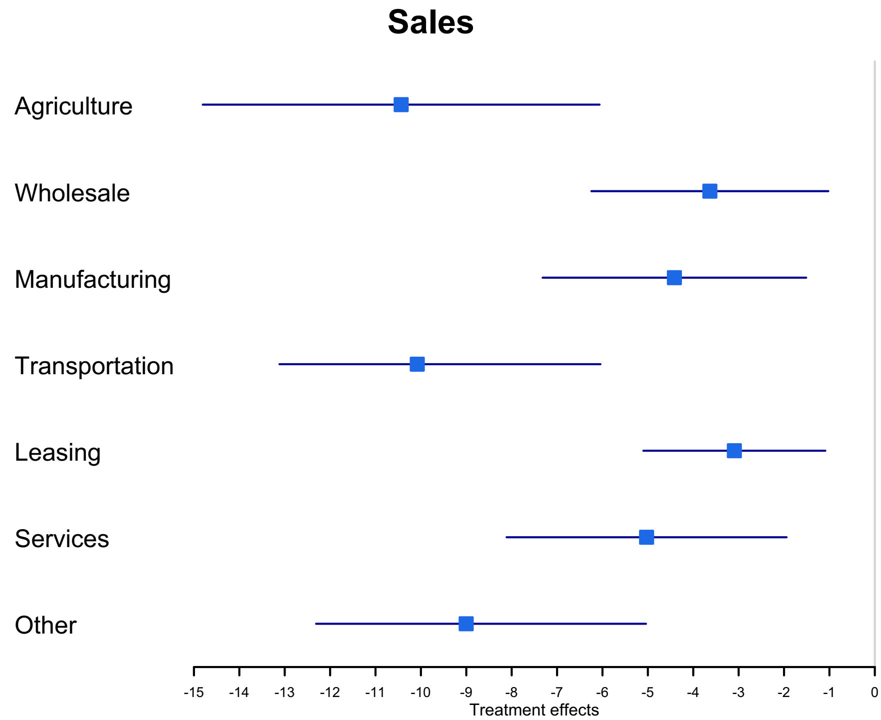

FIG. 4. Heterogeneity analysis of sales based on industry.

together, the above coping strategies can have a positive impact on firm performance on both the demand side and the supply side, and hypothesis 4 of this paper is valid. In addition, the coefficient curves show a rapid downward trend in the initial period. The reasonable explanation lies in the fact that adopting an adaptive strategy significantly mitigates the negative effect of drought shocks compared to not adopting any adaptive strategy. Therefore, the magnitude of the drought shock coefficient changes significantly in the initial stage. As the adaptive strategy is adopted, its marginal benefits gradually diminish. As a result, subsequent changes in the coefficients gradually flatten out.

7. Conclusions

This paper examines the impact of drought shocks on firm performance based on a generalized random forest model using 2015–19 CHFS data. The results find that drought shocks have a significant negative impact on firms' revenues, employment size, and firms' profits. The above results still hold significantly after some series of robustness tests. Furthermore, we demonstrate that the data used in the study contain enough information that affects firms' growth. Therefore, our model can effectively overcome the effects of omitted variables and other confounding factors on the results, and the strength of the data-driven approach is tested in the paper. The results of the mechanism analysis show that drought shocks affect enterprise performance mainly through the market demand effect and the cost effect. We also find that drought shocks have significant heterogeneous effects on firms with different asset sizes, industries, and ages. Further, we examine the effectiveness of different coping strategies in mitigating the negative impacts of drought shocks on firms. The study finds that improving regional agricultural infrastructure, promoting regional agricultural mechanization, and enhancing regional digital financial inclusion can effectively mitigate the negative impacts of drought shocks on firms' performance from both the demand side and the supply side.

The above findings, on the one hand, provide an effective reference for further clarifying and responding to the impacts of climate change on the economic development, and on the other hand, the findings provide new solutions for causal inference using machine learning methods in existing studies. Specifically, when considering the impact of climate change on the economy and society, the government and relevant departments should not only consider the agricultural sector but also its impact on enterprises deserves further attention. Particularly, micro, small, and medium-sized enterprises, which are the main carriers of employment in developing countries such as China and the economic backbone of these nations. Second, it is necessary for the government to include the firm sector and the agricultural sector in the same framework when formulating policies to address climate change. Specifically, governments can design an organizational structure that covers agriculture, industry, energy, transportation, construction, and other sectors; develop various measures including incentives, subsidies, and taxes; and provide training and guidance on technology and risk management to farmers and firm to promote regional actions to save energy, reduce emissions, and enhance climate resilience.

FIG. 5. Analysis of the effectiveness of adaptive strategies.

In addition, the government can also promote the participation of farmers, enterprises, and the public in actions to address climate change through the establishment of community outreach and the organization of information dissemination meetings and other publicity and education activities. From the perspective of China's current stage of development, improving agricultural infrastructure, upgrading the level of agricultural mechanization, and developing regional digital financial inclusion are effective measures to effectively deal with extreme weather caused by climate change. Specifically, we can develop digital financial inclusion by vigorously developing financial technology. At the same time, we can enhance regional mechanization through the development of agricultural mechanization services and so on.

In the experience of using machine learning methods for causal inference, our research has found that when the data used for a study contain enough information, the available data information can fit the structure of the unobserved features. Therefore, the results of the research are less likely to be affected by omitted variables and other confounding factors, and the advantages of the data-driven approach are highlighted. In addition, GRFs can effectively relax the assumption of performing causal inference and obtain robust results. Overall, the methodology used in the study has outstanding advantages in obtaining heterogeneous causal effects and dealing with omitted variables, and it is an effective and tractable tool for existing methods of estimating causal effects.

Acknowledgments. This study was supported by the MOE Project of Key Research Institute of Humanities and Social Sciences (22JJD910002). The authors declared no potential

conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement. The underlying data used in this study are available at CHFS (https://chfser.swufe.edu.cn/datas/) and China Urban Statistical Yearbook (https://ceidata.cei.cn/ and https://github.com/LijinLiu19/climate-change-and-firm-performance).

REFERENCES

Afridi, F., M. Bishnu, and K. Mahajan, 2023: Gender and mechanization: Evidence from Indian agriculture. *Amer. J. Agric. Econ.*, **105**, 52–75, https://doi.org/10.1111/ajae.12315.

Agamile, P., R. Dimova, and J. Golan, 2021: Crop choice, drought and gender: New insights from smallholders' response to weather shocks in rural Uganda. *J. Agric. Econ.*, **72**, 829–856, https://doi.org/10.1111/1477-9552.12427.

Amare, M., B. Shiferaw, H. Takeshima, and G. Mavrotas, 2021: Variability in agricultural productivity and rural household consumption inequality: Evidence from Nigeria and Uganda. *Agric. Econ.*, **52**, 19–36, https://doi.org/10.1111/agec.12604.

Arrighi, C., A. Carraresi, and F. Castelli, 2022: Resilience of art cities to flood risk: A quantitative model based on depth-idleness correlation. *J. Flood Risk Manage.*, **15**, e12794, https://doi.org/10.1111/jfr3.12794.

Athey, S., and S. Wager, 2019: Estimating treatment effects with causal forests: An application. *Obs. Stud.*, **5**, 37–51, https://doi.org/10.1353/obs.2019.0001.

——, J. Tibshirani, and S. Wager, 2019: Generalized random forests. *Ann. Stat.*, **47**, 1148–1178, https://doi.org/10.1214/18-AOS1709.

Basker, E., and J. Miranda, 2018: Taken by storm: Business financing and survival in the aftermath of Hurricane Katrina. *J. Econ. Geogr.*, **18**, 1285–1313, https://doi.org/10.1093/jeg/lbx023.

Bennett, A., and N. Kallus, 2019: Policy evaluation with latent confounders via optimal balance. arXiv, 1908.01920v1, https://doi.org/10.48550/arXiv.1908.01920.

Boudreaux, C. J., A. Jha, and M. Escaleras, 2023: Natural disasters, entrepreneurship activity, and the moderating role of country governance. *Small Bus. Econ.*, **60**, 1483–1508, https://doi.org/10.1007/s11187-022-00657-y.

Bozzola, M., and M. Smale, 2020: The welfare effects of crop biodiversity as an adaptation to climate shocks in Kenya. *World Dev.*, **135**, 105065, https://doi.org/10.1016/j.worlddev.2020.105065.

Branco, D., and J. Féres, 2021: Weather shocks and labor allocation: Evidence from rural Brazil. *Amer. J. Agric. Econ.*, **103**, 1359–1377, https://doi.org/10.1111/ajae.12171.

Chavas, J.-P., G. Rivieccio, S. Di Falco, G. De Luca, and F. Capitanio, 2022: Agricultural diversification, productivity, and food security across time and space. *Agric. Econ.*, **53**, 41–58, https://doi.org/10.1111/agec.12742.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins, 2018: Double/debiased machine learning for treatment and structural parameters. *Econometrics J.*, **21**, C1–C68, https://doi.org/10.1111/ectj.12097.

——, M. Demirer, E. Duflo, and I. Fernández-Val, 2025: Fisher–Schultz lecture: Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. *Econometrica*, **93**, 1121–1164, https://doi.org/10.3982/ECTA19303.

Clò, S., F. David, and S. Segoni, 2024: The impact of hydrogeological events on firms: Evidence from Italy. *J. Environ. Econ. Manage.*, **124**, 102942, https://doi.org/10.1016/j.jeem.2024.102942.

Coelli, F., and P. L. Manasse, 2014: The impact of floods on firms' performance. Quaderni—Working Paper DSE No. 946, 40 pp., https://doi.org/10.6092/unibo/amsacta/4025.

Conley, T., and U. Christopher, 2001: Social learning through networks: The adoption of new agricultural technologies in Ghana. *Amer. J. Agri. Econ.*, **83**, 668–673, https://doi.org/10.1111/0002-9092.00188.

Cui, X., and Z. Zhong, 2024: Climate change, cropland adjustments, and food security: Evidence from China. *J. Dev. Econ.*, **167**, 103245, https://doi.org/10.1016/j.jdeveco.2023.103245.

David, B., O. Victor, A. T. Justice, D. Emmanuel, and O. A. Bright, 2021: Rainfall shocks and household welfare: Evidence from northern Ghana. *Agric. Syst.*, **194**, 103267, https://doi.org/10.1016/j.agsy.2021.103267.

FAO, 2023: World Food and Agriculture – Statistical Yearbook 2023. United Nations, 367 pp., https://doi.org/10.4060/cc8166en.

Farzana, A., M. Kanika, and S. Nikita, 2022: The gendered effects of droughts: Production shocks and labor response in agriculture. *Labour Econ.*, **78**, 102227, https://doi.org/10.1016/j.labeco.2022.102227.

Feng, S., X. Wu, Z. Hao, Y. Hao, X. Zhang, and F. Hao, 2020: A database for characteristics and variations of global compound dry and hot events. *Wea. Climate Extremes*, **30**, 100299, https://doi.org/10.1016/j.wace.2020.100299.

Gannon, K. E., E. Castellano, S. Eskander, D. Agol, M. Diop, D. Conway, and E. Sprout, 2022: The triple differential vulnerability of female entrepreneurs to climate risk in sub-Saharan Africa: Gendered barriers and enablers to private sector adaptation. *Wiley Interdiscip. Rev.: Climate Change*, **13**, e793, https://doi.org/10.1002/wcc.793.

Grabrucker, K., and M. Grimm, 2021: Is there a rainbow after the rain? How do agricultural shocks affect non-farm enterprises? Evidence from Thailand. *Amer. J. Agri. Econ.*, **103**, 1612–1636, https://doi.org/10.1111/ajae.12174.

Hamilton, S. F., T. J. Richards, A. P. Shafran, and K. N. Vasilaky, 2022: Farm labor productivity and the impact of mechanization. *Amer. J. Agri. Econ.*, **104**, 1435–1459, https://doi.org/10.1111/ajae.12273.

He, C., Y. Zhang, A. Schneider, R. Chen, Y. Zhang, W. Ma, P. L. Kinney, and H. Kan, 2022: The inequality labor loss risk from future urban warming and adaptation strategies. *Nat. Commun.*, **13**, 3847, https://doi.org/10.1038/s41467-022-31145-2.

Holland, P. W., 1986: Statistics and causal inference. *J. Amer. Stat. Assoc.*, **81**, 945–960, https://doi.org/10.1080/01621459.1986.10478354.

IPCC, 2022: *Climate Change 2022: Impacts, Adaptation and Vulnerability.* Cambridge University Press, 2952 pp.

Jarrett, U., and Y. Tackie, 2024: Re-examining the effect of heat and water stress on agricultural output growth: How is sub-Saharan Africa different? *Agric. Econ.*, **55**, 515–530, https://doi.org/10.1111/agec.12830.

Khalili, N., M. Arshad, H. Kächele, Z. Farajzadeh, and K. Müller, 2021: Drought shocks and farm household consumption behaviour: Insights from Fars province of Iran. *Int. J. Disaster Risk Reduct.*, **66**, 102625, https://doi.org/10.1016/j.ijdrr.2021.102625.

Kirchberger, M., 2017: Natural disasters and labor markets. *J. Dev. Econ.*, **125**, 40–58, https://doi.org/10.1016/j.jdeveco.2016.11.002.

Kühl Teles, V., and C. Cesar Mussolini, 2012: Infrastructure and productivity in Latin America: Is there a relationship in the long run? *J. Econ. Stud.*, **39**, 44–62, https://doi.org/10.1108/01443581211192107.

Lazzaroni, S., and P. A. G. van Bergeijk, 2014: Natural disasters' impact, factors of resilience and development: A meta-analysis of the macroeconomic literature. *Ecol. Econ.*, **107**, 333–346, https://doi.org/10.1016/j.ecolecon.2014.08.015.

Leiter, A. M., H. Oberhofer, and P. A. Raschky, 2009: Creative disasters? Flooding effects on capital, labour and productivity within European firms. *Environ. Resour. Econ.*, **43**, 333–350, https://doi.org/10.1007/s10640-009-9273-9.

Liu, L., and L. Guo, 2023: Digital financial inclusion, income inequality, and vulnerability to relative poverty. *Soc. Indic. Res.*, **170**, 1155–1181, https://doi.org/10.1007/s11205-023-03245-z.

——, and Y. Wu, 2024: Drought shocks, adaptive strategies, and vulnerability to relative poverty. *Nat. Hazards*, **120**, 12 679–12 703, https://doi.org/10.1007/s11069-024-06704-z.

Noy, I., and T. B. Vu, 2010: The economics of natural disasters in a developing country: The case of Vietnam. *J. Asian Econ.*, **21**, 345–354, https://doi.org/10.1016/j.asieco.2010.03.002.

Okubo, T., and E. Strobl, 2021: Natural disasters, firm survival, and growth: Evidence from the Ise Bay typhoon, Japan. *J Regional Sci.*, **61**, 944–970, https://doi.org/10.1111/jors.12523.

Pan, X., and B. Qiu, 2022: The impact of flooding on firm performance and economic growth. *PLOS ONE*, **17**, e0271309, https://doi.org/10.1371/journal.pone.0271309.

Paudel, J., and H. Ryu, 2018: Natural disasters and human capital: The case of Nepal's earthquake. *World Dev.*, **111**, 1–12, https://doi.org/10.1016/j.worlddev.2018.06.019.

Quartey, P., E. Turkson, J. Y. Abor, and A. M. Iddrisu, 2017: Financing the growth of SMEs in Africa: What are the contraints to SME financing within ECOWAS? *Rev. Dev. Finance*, **7**, 18–28, https://doi.org/10.1016/j.rdf.2017.03.001.

Rubin, D. B., 1974: Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.*, **66**, 688–701, https://doi.org/10.1037/h0037350.

Shibia, A. G., 2024: Firms' use of formal and informal finance in coping with droughts and floods: Experiences from Kenya. *Climate Dev.*, **16**, 9–24, https://doi.org/10.1080/17565529.2023.2176187.

Skidmore, M., and H. Toya, 2002: Do natural disasters promote long-run growth? *Econ. Inq.*, **40**, 664–687, https://doi.org/10.1093/ei/40.4.664.

Straffelini, E., and P. Tarolli, 2023: Climate change-induced aridity is affecting agriculture in northeast Italy. *Agric. Syst.*, **208**, 103647, https://doi.org/10.1016/j.agsy.2023.103647.

Wager, S., and S. Athey, 2018: Estimation and inference of heterogeneous treatment effects using random forests. *J. Amer. Stat. Assoc.*, **113**, 1228–1242, https://doi.org/10.1080/01621459.2017.1319839.

Wang, Y., and D. M. Blei, 2019: The blessings of multiple causes. *J. Amer. Stat. Assoc.*, **114**, 1574–1596, https://doi.org/10.1080/01621459.2019.1686987.

Zhou, F., and W. Botzen, 2021: Firm level evidence of disaster impacts on growth in Vietnam. *Environ. Resour. Econ.*, **79**, 277–322, https://doi.org/10.1007/s10640-021-00562-0.