Living standards shape individual attitudes on genetically modified food around the world

Sebastian Levi (Hertie School)

2021-05-05

Please note: This is a working paper. Comments and suggestions are very welcome:

levi@hertie-school.org

Agricultural biotechnology can help to sustainably intensify food production, but negative public opinion hinders the deployment of genetically modified crops and livestock. Previous research shows negative consumer attitudes in the Global North to be primarily driven by limited trust and religiosity, but public opinion in the Global South remains largely unexplored. Here, analyzing individual attitudes across 142 countries with a random forest model, I show that people in low-income countries are significantly more positive towards genetically modified food than those living in high-income countries. Globally, individual attitudes are primarily determined by living standard, agricultural output, and prevalence of undernourishment. Country income levels also moderate how demographic characteristics predict attitudes on bioengineered food. Highly educated urban men are most optimistic about agricultural biotechnology in high-income countries, while women, individuals living in rural areas, and those with little education are the most hopeful demographic in low-income countries. These results indicate that individual views are largely determined by the societal benefits expected from agricultural biotechnology and suggest that the conditions for further deregulation of genetically modified food are most favorable in low-income countries.

Keywords: genetically modified crops, genetically engineered food, public opinion, consumer attitudes, Global South, machine learning

1. Introduction

More than two billion people have experienced moderate or severe levels of food insecurity in 2019, a number that is likely to rise given current and projected pressures on the global food

system resulting from population growth and climate change (Bebber et al., 2013; FAO, 2020; Myers et al., 2017; Springmann et al., 2016). Genetically improving crops and livestock promise to increase the resilience of food production and may thereby help to achieve global and sustainable food security (Steinwand & Ronald, 2020). Over the last years, genetically modified (GM) crops were cultivated in a growing number of countries, leading to sizable gains in yield, reduced herbicide usage, and increased farmer income, especially in the Global South (Hong et al., 2021; ISAAA, 2020; Paul et al., 2018). Meanwhile, research shows that bioengineered livestock may increase food production, advance human and animal welfare, and reduce the environmental impact of farming (Carlson et al., 2016; McFarlane et al., 2019; Tait-Burkard et al., 2018). Moreover, recent advances in precise genome editing allow to further enhance the efficiency of genetic modifications while simultaneously limiting the risks from off-target mutations (Woo et al., 2015; Zhang et al., 2020).

Despite these prospects for global food security and the advancements in biosafety, public opinion on genetically engineered food remains largely skeptical, driven by concerns about food safety, suspicions that genetic engineering would exclusively benefit international companies, and unease with the peculiar unnaturalness of GM products (Blancke et al., 2015; Gaskell et al., 2000; Kato-Nitta et al., 2019; Kennedy & Thigpen, 2020; Marris, 2001). Existing analyses of attitudes on genetically modified organisms (GMO) in the US, Europe, and other high-income countries find public opinion to be primarily driven by gender, education, religiosity, and trust in regulatory agencies (Akin et al., 2017; Bonny, 2003; Durant & Legge, 2005; Gaskell et al., 2000; Legge Jr. & Durant, 2010). In contrast, the determinants of attitudes in Latin America, Africa, and Asia, regions where the cultivation of GM crops has recently expanded most rapidly, remain largely unexplored.

To fill this gap, I analyze data from the World Risk Poll, which collected attitudes on genetically modified food from more than 150,000 individuals in 2019 (Lloyd's Register Foundation & Gallup, 2020). The data is nationally representative to 142 countries and territories that jointly cover more than 95 percent of the world population (see Methods). I use a random forest model, a versatile method based on machine learning, to estimate the degree in which individual- and country-level conditions predict individual attitudes on genetically modified food. I further apply ancillary random forest techniques to analyze the functional form of these predictions, local

predictor effects across countries, and the degree to which country income levels moderate the magnitude and direction in which demographic traits predict individual attitudes (see Methods for details). Throughout this article, I use the terms genetic engineering and genetic modification interchangeably, encompassing both cisgenic and transgenic modifications, consistent with the USDA glossary (U.S. Department of Agriculture, 2021). Existing studies indicate that most people are unaware of recent innovation in precise genome editing (Bundesinstitut für Risikobewertung, 2017), that attitudes do not significantly change depending on whether the terms gene editing or gene modification are used (Weisberg et al., 2017), and that even after the differences between gene editing and gene modification are explained, attitudes remain largely similar for both technologies compared to attitudes on conventional breeding methods (Kato-Nitta et al., 2019; Shew et al., 2018).

2. Materials and Methods

2.1. Data Collection and Cleaning

2.1.1. Data Collection

This article draws on data from the World Risk Poll (WRP), a global survey conducted in 2019 by Gallup, a polling company (Lloyd's Register Foundation & Gallup, 2020). The WRP collected data from more than 150,000 individuals and is nationally representative to the adult population of 142 countries and territories (Lloyds Register Foundation & Gallup, 2020). Please see Supplementary Data 1 for sample sizes in each country. The WRP includes information from respondents that are typically not featured in other surveys, such as people living in remote areas, illiterate people without access to telephone and internet, and individuals living in countries with active conflicts. For countries in which the interviews were conducted face-to-face, sampling units were stratified by population size or geography with clusters at one or multiple stages. Depending on the availability of population information, sample selection was based on probabilities proportional to population size, otherwise, simple random selection was used. For countries in which interviews were conducted by phone, Gallup used random-digit-dialing or a nationally representative list of phone numbers. Post-stratification weights were constructed

based on gender, age, and where reliable data were available, education or socioeconomic status (Lloyds Register Foundation & Gallup, 2020).

2.1.2. Operationalization

I operationalize the dependent variable, attitude on genetically modified food, with the survey question L4A from the WRP. The item is phrased “Please tell me whether you think each of the following will mostly help or mostly harm people in this country in the next 20 years.

Genetically modified food” and offers the response options “mostly help”, “mostly harm”, and “Don’t have opinion”. The volunteer-only options “Neither”, “Don’t know”, and Refusal were also recorded.

All individual-level explanatory variables were composed based on data from the WRP such as age, gender, education, income quintile, number of children in household, and urbanicity. The variable “Support government regulation” is based on the survey item “In general, do you think the government should require businesses to adopt safety procedures and rules, or not?” (response options “Yes” or “No”). The variable “Trust government regulation” is based on the item “In general, in your opinion does the government do a good job ensuring that the following are safe, or not? The food you buy” (response options “Yes” or “No”). The variable “Internet usage” is based on the item “Have you used the Internet, including social media, in the past 30 days?” (response options “Yes” or “No”). The variable “General risk perception” is based on the item “When you hear the word risk, do you think more about opportunity or danger?” (response options “Opportunity” or “Danger” but volunteered options “Both” and “Neither” were also recorded). The variable “Statistical literacy” is based on the survey item “Do you think that 10% is bigger than 1 out of 10, smaller than 1 out of 10, or the same as 1 out of 10? If you do not know, please just say so.” Respondents answering that 10% is the same as 1 out of ten were coded as statistically literate, all others as not. The variable “Experience food poisoning” is based on the item “Have you or someone you personally know, experienced serious harm from any of the following things in the past two years? Eating food” (“Yes” or “No”). The variables “Trust Neighbor” and “Trust stranger” are based on the items “Suppose you lost a small bag that contained items of great financial value to you that had your name and address written on it. If it were found by each of the following people, in general, how likely is it that it would be returned

to you with all of its contents? A stranger / A Neighbor” (response options “Very likely”, “Somewhat likely”, or “Not likely at all”). The variable “Most trusted source” is based on the item “Considering the sources of information you would access, which one would you trust most to provide information about food safety?”. Response options are “Friends and Family”, “Medical professionals such as your local doctor or nurse”, “Newspaper, television, or radio” (relabeled as “Media”), “Internet / social media”, the country-specific relevant government agency responsible for providing food safety information, “the packaging or label of the food”, “a famous person you like”, “local religious leaders”, “would trust none of them”, “Some other source”. The variables “Trust information from Friends / Government / Religious Leaders / Food Labels / Internet” are based on the items “Suppose you wanted to find out if the food you eat is safe. Would you look to any of the following sources for information, or not? Friends and Family / country-specific government agency responsible for food safety / Local religious leaders/ The packaging or label on the food/ The internet or social media”. The response options for each of these items were “Yes” or “No”. For all items, volunteered “Don’t know” responses and refusals to answer the question were also recorded.

I supplement the individual-level information with country-level data. In particular, I use data from the World Bank for living standard (in GDP per Capita, purchasing power parity) (World Bank, 2021b), agricultural output (in % of GDP) (World Bank, 2021a), and scientific output (in scientific and technical articles published in the country) (World Bank, 2021c), from the Food and Agriculture Organization (FAO) for prevalence of undernourishment (in % of the population) (Food and Agriculture Organization of the United Nations, 2021), and from Freedom House for civil liberty (measured with an index) (Freedom House, 2021). The variable Good Governance is a composite indicator based on the averaged Worldwide Governance Indicators Regulatory Quality, Control of Corruption, Political Stability, Rule of Law, and Government Effectiveness (World Bank, 2021d). For all these country-level variables, a three-years lagged average (2017-2019) was used to allow a sufficient time lag for societal conditions to materialize on the individual-level. Please see Supplementary Table 2-3 for details and descriptive statistics of all explanatory variables considered in this article.

2.1.3. Data Pre-Processing

Prior to the descriptive and explanatory analyses, I removed all respondents from the sample who refused to disclose their attitude on genetically modified food (277 from originally 151,195 respondents). For individual-level variables, I recoded refusals to answer a survey question as missing. Moreover, Don't know (DK) answers to the question of formal education was also recoded as missing. For all other questions, DK answers were recoded as substantive category.

Afterwards, I supplemented missing country-level information with ancillary data. I used external data for living standard in Taiwan and Kosovo and for agricultural output in Taiwan, Turkmenistan, Libya, Venezuela, and Canada. Moreover, I used ancillary data on undernourishment in the Palestine Territories, Tajikistan, Laos, Niger, Yemen, Zambia, Guinea, Moldova, Zimbabwe, Lydia, Kosovo, Singapore, and Uganda. Please see Supplementary Table 4 for details on the estimates collected from external sources. Remaining missing data were imputed, including data on undernourishment in Bahrain and the number of scientific articles published in Taiwan and Hong Kong. The variables with the highest frequency of missing observations were the “Most trusted source” (missing for 7,241 observations, representing 4.7% of the dataset) and “Trust information from religious leaders” (missing for 3,886 observation, equivalent to 2.5% of all individuals in the dataset). Please see Supplementary Table 5 for further details on the number of missing observations for each explanatory variable.

I decided to impute missing observations, as case-wise deleting these observations from the sample would have distorted the carefully assembled and weighted sample and hence corrupt its representativeness. Moreover, case-wise deletion or mean imputation of missing values may introduce statistical error and bias, if the missingness is not distributed purely at random. Hence, I use a regression-based multiple imputation technique which is the safest way to address missing values while mitigating the risk of statistical bias (Kang, 2013). I specifically use the *missRanger* package, which is a computationally optimized version of the *missForest* algorithm (Mayer, 2019). The *missForest* algorithm is an iterative random forest imputation technique that is able to impute moderate to high amounts of missing data, even under presence of mixed type data and strongly interacting covariates (Stekhoven & Buhlmann, 2012). The *missForest* mechanism has found to outperform all other currently available imputation techniques (Tang & Ishwaran, 2017).

For the random forest analysis explaining why individuals expect genetically modified food to either help or harm people living in their country, I excluded individuals who neither expect GM food to help or harm people in their country, people who do not know how to answer this question, and individuals who have no opinion on this matter. In total, 38,973 individuals were removed from the sample, leaving 114,945 respondents in the final dataset. As a robustness check, I also calculated an alternative model specification in which I left respondents with no opinion on genetically modified food in the dataset.

2.2. Descriptive Analysis

I estimate the global opinion on genetically modified food (displayed in Fig. 1a) based on data from the WRP using population-scaled survey weights to correct for differences in sampling probability, differences in sample sizes, and national population sizes. Please see Supplementary Table 1 for the numerical estimates and the respective standard errors. I also calculate national estimates of public opinion on genetically modified food using post-stratification weights (displayed in Fig. 1b). These country-specific estimates are provided in Supplementary Data 1. The article states that concerns about GM food are significantly higher in high-income countries compared to the rest of the world and that hopes about GM food are significantly higher in low-income countries compared to the rest of the world. Both claims are formally tested using a Wilcoxon rank sum test in which the means of believing that GM food mainly helps (or mainly harms) among individuals in low-income countries (high-income countries) are compared to the means of these variables among individuals from all other country income groups. The results for both tests confirm the hypothesis and are statistically significant at a 0.1% significance level.

2.3. Random Forest Analysis

2.3.1. Random Forest Model

The results in this article are estimated with a random forest model (Breiman, 2001; Strobl et al., 2009). Random forests are a versatile machine learning technique frequently used in genetics and neurosciences and increasingly also applied in social science analyses (Athey, 2019; Molina &

Garip, 2019). Their ability to consider large numbers of mixed-type data and to model non-linear relationships with high-dimensional interactions and complex data-structures makes them particularly suitable to analyze the diverse conditions that influence attitudes on GM food around the world (Karpiévitch et al., 2009; D. P. Martin, 2015; Strobl et al., 2009).

I use the *party* package in *R* to estimate the non-randomized partitioning tree models displayed in Figure 3 and Figure 6. The tree model in Figure 6c analyzes individual attitudes across all 27 member states of the European Union: Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, and Sweden.

For the main random forest model, I use the *randomForestSRC* package in *R*, which is particularly suitable to estimate random forest models in classification settings where both dependent and independent variables are categorical (Ishwaran & Kogalur, 2019). Here, splitting decisions are made by maximizing the purity in the class of the dependent variable in each daughter node (Kogalur & Ishwaran, 2020). The model's hyper-parameters were optimized to maximize the model's prediction accuracy using the auto-tuning function provided in the package. In particular, the number of predictors considered in each splitting decision and the forest's average terminal node size were automatically optimized, while the numbers of trees used to estimate the variable importance was set to 1000, a standard parameter. I also specify population-scaled survey weights as case weights to account for differences in sampling probability, sample sizes, and population sizes.

The estimated variable importance displayed in Figure 2 is calculated using unscaled mean contribution to prediction accuracy (permutation importance). I use *randomForestSRC*'s computationally-optimized "random" algorithm for estimating variable importance. The unscaled permutation importance was found to biases prevalent in scaled importance measures or impurity importance measures (Strobl et al., 2007; Strobl & Zeileis, 2008). The permutation importance may be biased in the presence of strong multicollinearity, but the covariates considered in this model are only moderately correlated with each other. Please see Supplementary Table 6 for estimated correlation coefficients among all numerical covariates.

2.3.2. Ancillary Analyses

Random forest models tend to be more difficult to interpret than conventional regression analyses, which makes it challenging to use them to “explain” outcomes (as opposed to predict them) (Elshawi et al., 2019). To increase the interpretability of the random forest model output, I use three ancillary knowledge extraction techniques.

First, I conduct a permutation-based significance test, which is a (two-sided) significance test that estimates the likelihood that the variable importance results purely from random chance. In this test, the estimated variable importance is compared to the distribution of variable importance under the Null hypothesis of no association, which is generated by fitting the model a high number of times while randomly permutating the response variable (Ojala & Garriga, 2010). I have calculated the distribution of the Null hypothesis and the significance levels by calculating 1000 permuted model iterations using the package *rfPermute* (Archer, 2019; Knijnenburg et al., 2009). Please see Supplementary Table 7 for numerical representations of the estimated significance levels of all predictors.

Second, I analyze the overall form and direction in which each explanatory variable relates to the outcome variable using a technique called partial variable predictions. This method allows to examine the form in which an explanatory variable is related to an outcome variable by calculating n predictions of the outcome variable on n evenly spaced-out values of an explanatory variable while integrating out the effect of all other co-variates at this position (Ehrlinger, 2015; Ishwaran et al., 2010). I calculate the partial variable predictions with the *randomForestSRC* package (Ishwaran & Kogalur, 2019). Please see Supplementary Figures 1-3 for visual representations of the partial variable predictions for all 27 explanatory variables.

Third, I estimate the local predictor importance for each single observation and aggregate this local importance to the national level. This metric is used to create Figure 5. The local predictor importance is a metric of random forests that measures the average times a specific observation is misclassified, given a certain explanatory variable is randomly permuted (Breiman, 2001). I calculate the local importance with the *randomForest* package (Breiman, 2001). Please see

Supplementary Data 2 for the local predictor importance of all 27 explanatory variables aggregated to all 142 countries.

2.3.3. Robustness Tests

As a robustness test to the central model presented and discussed in this article, I estimate a model specification in which all respondents who have no opinion on genetically modified food are left in the sample. Please see Supplementary Table 8 for the results of this robustness test. In this alternative specification, the main results remain robust, as all predictors remain statistically significant. Moreover, the magnitude of the variable importance remains highly robust for most predictors, in particular living standard, agricultural output, prevalence of undernourishment, and scientific output. The importance of several individual-level variables has significantly increased in the robustness specification, mostly among variables that contain a Don't Know response option. The increased variable importance for those variables may be attributed to an increased value of these variables for model's ability to differentiate individuals with an opinion on genetically modified food from individuals who have no opinion on genetically modified food.

Figure 1 **a** Population-weighted estimates of share of global population believing that genetically modified food will mostly help or harm people living in their country **b** estimated share of the national population believing that GM food will mostly help people living their country.

3. Results

3.1. Globally, most people expect genetically modified food to do more harm than good

Approximately half of the world's population believes that genetically modified food will mostly harm people living in their country. Only one of five individuals is convinced that agricultural biotechnology will mostly bring benefits. Moreover, one of three individuals does not know enough about the topic or does not have an opinion (Figure 1a). Public opinion on GM food varies substantially across countries (Figure 2b). In 85 out of 142 countries, most of the national adult population is explicitly opposed to GM food. Public skepticism is highest in Central Asia, especially in Turkmenistan and Georgia where 91 percent and 81 percent of the population believes that GM food will mostly harm people in their country. Public opinion is most positive in Sub-Saharan Africa, particularly in Eswatini and Ethiopia where 87 percent and 73 percent of the population believes that GM food will mostly help people living in their country. Only in four countries (Nigeria, Madagascar, Ethiopia, and Eswatini) most of the population is explicitly optimistic about the prospects of agricultural biotechnology.

In general, people in high-income countries are significantly more pessimistic about genetically modified food than those living in other world regions. Moreover, people living in low-income countries tend to be more optimistic about GM food compared to individuals living elsewhere (see Methods). Please see Supplementary Data 1 for country-level estimates of public opinion for all 142 countries.

3.2. Living standard, agricultural output, and undernourishment predict attitudes on GM food

I use a random forest model to estimate the probability that a given individual expects GM food to either help or harm people living in their country. The model considers 27 explanatory variables that were identified in previous research as potentially influencing attitudes on GM food or GMO more broadly and for which sufficient data were available. Figure 2 displays the estimated variable importance of all predictors, measured in mean increase in prediction accuracy (see Methods for details). The model results demonstrate that individual attitudes on genetically modified food are best predicted by country-level living standard (in GDP per Capita

in purchasing power parity), agricultural output (in share of national economic output), and prevalence of undernourishment (in share of national population). Further conditions that substantially increase the accuracy of correctly predicting attitudes on GM food are country-level civil liberty, country-level scientific output, individual support of government regulation and age. Individual internet usage, risk perception, trust in information from local religious leaders, country-level governance quality, trust in government regulation, formal education, and trust in strangers are moderately predictive of attitudes on GM food. In contrast, gender, income quintile, and statistical literacy hardly increase the accuracy of predicting individual attitudes on genetically modified food across 142 countries.

Figure 2 Estimated variable importance for predicting attitudes on genetically modified food. Variable importance is measured as independent contribution to prediction accuracy. Country-level conditions are denoted with “c”. The figure shows that country-level living standard,

agricultural output, and undernourishments are most important for correctly predicting individual attitudes on genetically modified food across 142 countries.

Despite the low variable importance of many predictors, most of them increase the accuracy of predicting GMO attitude at a one percent statistical significance level. However, statistical literacy, trust in information from food labels, and trust in information from the internet are only significant at a five percent level. The number of children in the household and personal experience of food poisoning are only significant at a ten percent level. Please see the Methods and Supplementary Table 2 for a detailed discussion of all explanatory variables and Supplementary Table 3 for numerical model results.

Figure 3 Global partitioning tree model on attitudes about genetically modified food (n=114,945). The grey boxes indicate the variables on which splitting decisions are made, while the dark green boxes indicate the significance level of these splitting decision. The terminal nodes display the share of individuals believing that GMO food either helps or harm people living in the country. The figure shows that the root node, Agricultural output, has a major influence on classifying observations.

Figure 3 displays a non-randomized recursive partitioning tree model to visualize the relationships of some of the most important variables with individual attitudes on GM food. The figure shows that women who do not trust government regulation and who live in countries with little agricultural output and low prevalence of undernourishment are globally most skeptical

about GM food (Figure 3; terminal node 6). Among individuals with these characteristics, almost 90 percent believe that GM food will mostly harm people living in their country. In contrast, individuals who live in countries with a large agricultural output, who have obtained either primary or tertiary education, and who trust both government regulation and information from local religious leaders are globally most optimistic about GM food (Figure 3; terminal node 9). Among these individuals, more than 60 percent believe that GM food will mostly help people living in their country.

Using partial variable predictions, I also estimate the form in which the model's predictors relate to individual attitudes on GM food (see Methods for details). Figure 4 displays the partial variable predictions of eight predictors that are either highly predictive in the model or frequently discussed in the literature on GMO attitudes. Panel 4a shows that the relationship between country-level living standard and individual attitudes on GM food is highly non-linear. Individuals are most likely to be optimistic about agricultural biotechnology if they live in countries with low living standards. The probability of expecting benefits from GM food declines by nine percent for increasing values of GDP per capita up until a living standard of approximately \$28,000 (equivalent to the living standard in Russia). A high standard of living beyond this value is again associated with an increasing likelihood of optimism about agricultural biotechnology.

Panel 4b shows a similar pattern for the prediction effect of country-level agricultural output. Individuals are least likely to hope for benefits from genetically engineered food if they live in countries with a moderate agricultural output. The probability of hoping that GM food mainly helps people increases both for individuals living in countries with a very small agricultural output (by four percent) as well as for individuals living in countries with a large agricultural output (by five percent). Panel 4c shows that individuals are eight percent more likely to expect benefits from agricultural biotechnology if they live in a country where more than 15 percent of the population suffer from undernourishment. Moreover, panel 4d visualizes the correlation of scientific output, measured in the total number of scientific articles published in a country, with the individual probability of hoping for benefits from agricultural biotechnology. Individuals living in countries with a strong scientific output are one percent more likely to expect benefits

from GM food compared to individuals living in countries with low or moderate scientific output.

Figure 4 Partial variable predictions of six predictors of attitudes on GM food. The panels display the estimated probability of hoping that GM food will mainly help people as opposed to believing that it will mainly harm people if the following variable change from their minimum to their maximum: **a** living standard, **b** agricultural output, **c** prevalence of undernourishment, **d** scientific output, **e** age, **f** education, **g** trust in information from religious leaders, and **h** gender. Variables displayed in panel a – d are country-level variables, while panels e-h display individual-level variables.

Figure panels 4e-f visualize the partial variable predictions of four individual-level variables that are frequently discussed in the literature on GMO attitudes. Panel 4e shows that young individuals are most likely to expect benefits from GM food (by 6 percent). Panel 4f show that individuals with secondary education are least convinced by the prospects of agricultural biotechnology. Both individuals with less formal education as well as individuals with at least tertiary education are estimated to be three and two percent more likely to be optimistic about

GM food, respectively. Finally, panels 4g shows that individuals who trust information from local religious leaders are more likely to believe in benefits of GM food, while panel 4h shows that men are almost four percent more likely to expect benefits from agricultural biotechnology compared to women. Please see Supplementary Figures 1-3 for the partial variable predictions of all 27 explanatory variables considered in this article.

3.3. Country income moderates the magnitude and direction of individual-level predictors

The model results further show that country income levels moderate the magnitude and direction in which individual-level conditions predict attitudes on GM food. Figure 5 visualizes how the prediction effect of five individual-level variables change depending on country and country income group. Panels 5a-f visualize the changes in prediction direction by displaying partial variable predictions stratified by country-income groups. Panels 5f-j show the cross-national variation in the prediction magnitude by displaying estimated local predictor importance aggregated on the country level (see Methods for details).

Panel 5a shows that individuals who trust governmental food regulations are 7-10 percent more likely to hope for benefits from agricultural biotechnology across all country income groups. However, the prediction effect of trust in government regulation is strongest in high-income countries. Panel 5f shows that trust in government regulation tends to be most predictive of individual attitudes in Mali, Morocco, and France. It is much less important in many countries in Sub-Saharan Africa and in China.

Panel 5b shows a large gender gap in GM food attitudes in high-income countries, where men are almost 9 percent more likely to be optimistic about GM food compared to women. However, this gender gap is marginal in all other country groups. In low-income countries, the relationship is even reversed, with men being insignificantly more skeptical about benefits from GM food compared to women. Panel 5g shows that the prediction effect of gender is strongest in the Great Britain, Australia, and Sweden. It is very small across Latin America, Africa, and most of Asia.

Figure 5 Variation of prediction effect across countries and country income groups. Panel **a-e** display the partial variable predictions of **a** trust in government regulation, **b** gender, **c** age, **d** education, and **e** urbanicity stratified by country income group. Panel **g-x** display the local predictor importance of **f** trust in government regulation, **g** gender, **h** age, **i** education, and **j** urbanicity aggregated at the country level.

Panel 5c visualizes the partial prediction effect of age on GM food attitudes stratified by country income group. The model results show that younger individuals are most likely to be optimistic about GM food across all country groups. The prediction effect is strongest in high-income countries, where the youngest individuals are up to nine percent more likely to hope for benefits from GM food compared to older people. Panel 5h shows that age is most predictive of GM food attitudes in South Korea, Finland, and Switzerland. Age tends to be much less important for correctly predicting GMO attitude across most of Latin America, Africa, and South Asia.

Panel 5d visualizes the prediction effect of education on GM food attitudes, stratified by country-income. Across all country groups, individuals with secondary education are least optimistic about genetically modified food. The prediction effect of education is strongest in low-income countries. Here, individuals with primary education are most likely to believe in the benefits of GM food – approximately six percent more than individuals with secondary education and five percent more than individuals with tertiary education. The prediction effect is much lower in middle- and high-income countries. Moreover, the direction in which education correlates with individual attitudes varies across country groups. In low- and middle-income countries, the least educated individuals tend to be most hopeful about GM food. In high-income countries, in contrast, the most educated individuals are the most optimistic about agricultural biotechnology. Panel 5i show that the local importance of education is strongest in Burkina Faso, Niger, and Madagascar.

Panel 5e visualizes the prediction effect of urbanicity on GM food attitudes. It shows that urbanicity is strongly correlated to GM food attitude in low-income and lower middle-income countries, but not in upper middle- and high-income countries. Moreover, the prediction direction varies across country income groups. In low- and middle-income countries, individuals in rural areas are substantially more likely to expect positive developments from GM food than urban residents. In high-income country, urban residents tend to be most optimistic about the benefits from agricultural biotechnology. Panel 5j visualizes the local predictor importance of urbanicity, showing that urbanicity strongly predicts GM food attitudes in Ethiopia, Niger, and Burkina Faso but not in the Americas, Australia, and most of Asia. Please see Supplementary Data 2 for the local predictor importance of all 27 variables and all 142 countries.

These extensive variations in prediction effects indicate that the conditions that shape public opinion on genetically modified food vary themselves substantially across different regions (see Discussion). Figure 6 exemplifies the diversity of the individual determinants of opinion on GM food by displaying recursive partitioning tree models for attitudes in the US, the European Union, Madagascar, and Ethiopia.

Figure 6 National recursive partitioning tree model on attitudes towards GM food. The panels display recursive partitioning tree models predicting attitudes on GM food among individuals having an opinion on GM food in **a** the United States (n=866), **b** Madagascar (n=930), **c** European Union (EU, n=21,662), and **d** Ethiopia (n=938). The grey boxes indicate the variables on which splitting decisions are made, while the dark green boxes indicate the significance level of the splitting decision. The terminal nodes display the share of respondents believing that GMO food either helps or harm people living in their country.

In the US and Europe, places with a high standard of living and insignificant rates of undernourishment, attitudes on genetically modified food are primarily determined by trust, education, and gender. In the US (Panel 6a), individuals are most likely to be optimistic about genetically modified food if they trust government regulation, if they consider government agencies or food labels to be the most trustworthy information source on food safety, and if they do not trust information from religious leaders. In contrast, US Americans who do not trust

government regulation and who have no tertiary education are the least likely to hope for benefits from GM food. In the European Union (EU; Panel 6c), individuals are most likely to hope for benefits from genetically modified food if they trust government regulation, if they generally feel safe, and if they consider either the internet, government agencies, famous people, or local religious leaders to be the most trustworthy source on food safety. In contrast, Europeans who neither trust governmental regulation nor information from governmental sources and who even do not trust their own neighbors tend to be least likely to believe that genetically modified food will mostly help people living in their country.

Attitudes on GM food in Madagascar and Ethiopia, two countries with a particularly positive public opinion on agricultural biotechnology, are not primarily determined by trust, but rather by urbanicity and education. In Madagascar (Panel 6b), individuals are most likely to hope for benefits from GM food if they have obtained little formal education, live in rural areas, and generally feel safe. Madagascans who are highly educated tend to be most skeptical about genetically modified food. In Ethiopia (Panel 6d), individuals are most likely to hope for benefits from genetically modified food if they live in rural areas, feel generally safer than in the past, and if they are older than 17 years. Urban Ethiopians are significantly more skeptical about agricultural biotechnology.

3. Discussion and Conclusion

This article shows that, globally, individual attitudes on genetically modified food are largely determined by country-level circumstances like living standard, agricultural output, and prevalence of undernourishment. These results indicate that the cross-national variation in public opinion about genetically engineered food can be attributed to the different societal benefits expected from agricultural biotechnology. Individuals living in countries with low standards of living, prevalence of undernourishment, and little agricultural output may be particularly optimistic about genetically modified food because they are at a higher risk of being food insecure and hope to increase crop resilience using biotechnology (Aerni & Bernauer, 2006; Ma et al., 2018; H. M. Martin et al., 2017). In contrast, individuals living in countries with a high standard of living, an extensive agricultural sector, and a large academic output may hope for

benefits from agricultural biotechnology because they are more likely than others to realize profits from agricultural biotechnology developed and patented in these countries.

The results further establish that insights derived from previous analyses of GMO attitude in the Global North cannot universally be generalized to the Global South given that the prediction effects of several individual conditions vary, and sometimes even reverse, across countries. Consistent with existing studies, this article finds younger and more trusting individuals to be more optimistic about genetically modified food across all country groups (Bonny, 2003; Diamond et al., 2020; Durant & Legge, 2005; Gaskell et al., 2000; Hunt & Frewer, 2001; Legge Jr. & Durant, 2010). However, both age and trust in government regulation are much more predictive of attitude in high-income than of attitudes in low-income countries. The prediction effect of gender, in contrast, varies strongly across country groups. While men are substantially more likely to be optimistic about agricultural biotechnology in high-income countries, this correlation is much weaker in middle-income countries, indicating that the proposed gender gap in GMO attitudes may only exist in the Global North (Elder et al., 2018). In low-income countries, the relationship is even reversed with women being marginally more likely to be hopeful about GM food compared to men. One interpretation of this heterogeneity may be that men in high-income countries are optimistic about GM food because they are more likely to work for biotechnology companies or to hold patents (McMillan, 2009). In low-income countries, women may be particularly hopeful about the prospects of agricultural biotechnology because they are more frequently exposed to food insecurity (World Food Programme, 2020).

Finally, this article also shows that education and urbanicity are much more important to correctly predict attitudes on GM food in low-income countries than in middle- or high-income countries. The prediction direction of both variables changes across country income groups. In high-income countries, the highest educated individuals and those living in urban areas are most optimistic about GM food, which is consistent with existing studies (Akin et al., 2017; Gaskell et al., 2000). However, this relationship is reversed in low-income countries. This heterogeneity can also be attributed to the different benefits expected from bioengineered food in the Global North and in the Global South. In high-income countries, profits from agricultural biotechnology are primarily realized by individuals working for biotechnology companies, who tend to be highly educated and based in urban areas. In low-income countries, genetically modified food

mostly benefits individuals working in agriculture, who tend to live in rural areas and frequently have obtained little formal education.

The observed cross-national differences in public opinion entail important implication for the regulation and deployment of genetically engineered food. The positive views on GM food in middle- and low-income countries already lead to an increasingly lenient regulation of GM crops and livestock in the Global South (Ledford, 2019; Menz et al., 2020). The disproportional benefits from agricultural biotechnology combined and the exceptionably favorable attitudes in low-income countries, especially in Sub-Sahara Africa, make further deregulations in these places even more likely (Malyska et al., 2016; Paarlberg, 2010). This will not come without complications, given the deficits in governance capacity in many low-income countries. Forthcoming regulations on genetically engineered food in these countries should be as efficient as possible to allow GM food to fulfill its promises on food security in the Global South, but also effective enough to maintain and increase trust in biosafety among citizens in the Global North and beyond.

4. References

- Aerni, P., & Bernauer, T. (2006). Stakeholder attitudes toward GMOs in the Philippines, Mexico, and South Africa: The issue of public trust. *World Development*, 34(2), 557–575.
- Akin, H., Rose, K. M., Scheufele, D. A., Simis-Wilkinson, M., Brossard, D., Xenos, M. A., & Corley, E. A. (2017). Mapping the Landscape of Public Attitudes on Synthetic Biology. *BioScience*, biw171. <https://doi.org/10.1093/biosci/biw171>
- Archer, E. (2019). *Package 'rfPermute.'* CRAN.
- Athey, S. (2019). The Impact of Machine Learning on Economics. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda* (pp. 507–547). University of Chicago Press.
- Bebber, D. P., Ramotowski, M. A. T., & Gurr, S. J. (2013). Crop pests and pathogens move polewards in a warming world. *Nature Climate Change*, 3(11), 985–988. <https://doi.org/10.1038/nclimate1990>
- Blancke, S., Van Breusegem, F., De Jaeger, G., Braeckman, J., & Van Montagu, M. (2015). Fatal attraction: The intuitive appeal of GMO opposition. *Trends in Plant Science*, 20(7), 414–418. <https://doi.org/10.1016/j.tplants.2015.03.011>

- Bonny, S. (2003). Why are most Europeans opposed to GMOs? Factors explaining rejection in France and Europe. *Electronic Journal of Biotechnology*, 6(1).
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Bundesinstitut für Risikobewertung. (2017). *Durchführung von Fokusgruppen zur Wahrnehmung des Genome Editings (CRISPR/Cas9)—Abschlussbericht*. <https://mobil.bfr.bund.de/cm/350/durchfuehrung-von-fokusgruppen-zur-wahrnehmung-des-genome-editings-crispr-cas9.pdf#search>
- Carlson, D. F., Lancto, C. A., Zang, B., Kim, E.-S., Walton, M., Oldeschulte, D., Seabury, C., Sonstegard, T. S., & Fahrenkrug, S. C. (2016). Production of hornless dairy cattle from genome-edited cell lines. *Nature Biotechnology*, 34(5), 479–481. <https://doi.org/10.1038/nbt.3560>
- Diamond, E., Bernauer, T., & Mayer, F. (2020). Does providing scientific information affect climate change and GMO policy preferences of the mass public? Insights from survey experiments in Germany and the United States. *Environmental Politics*, 29(7), 1199–1218. <https://doi.org/10.1080/09644016.2020.1740547>
- Durant, R. F., & Legge, J. S. (2005). Public Opinion, Risk Perceptions, and Genetically Modified Food Regulatory Policy: Reassessing the Calculus of Dissent among European Citizens. *European Union Politics*, 6(2), 181–200. <https://doi.org/10.1177/1465116505051982>
- Ehrlinger, J. (2015, January). *ggRandomForests: Visually Exploring a Random Forest for Regression*. <http://arxiv.org/abs/1501.07196>
- Elder, L., Greene, S., & Lizotte, M. K. (2018). The gender gap on public opinion towards genetically modified foods. *The Social Science Journal*, 55(4), 500–509. <https://doi.org/10.1016/j.soscij.2018.02.015>
- Elshawi, R., Al-Mallah, M. H., & Sakr, S. (2019). On the interpretability of machine learning-based model for predicting hypertension. *BMC Medical Informatics and Decision Making*, 19(1), 146–146. <https://doi.org/10.1186/s12911-019-0874-0>
- FAO. (2020). *The State of Food Security and Nutrition in the World 2020*. FAO, IFAD, UNICEF, WFP and WHO. <https://doi.org/10.4060/ca9692en>
- Food and Agriculture Organization of the United Nations. (2021). *Prevalence of undernourishment (percent) (3-year average)*. Suite of Food Security Indivators. <http://www.fao.org/faostat/en/#data/FS>
- Freedom House. (2021). *Country and Territory Ratings and Statuses, 1973-2019*.
- Gaskell, G., Allum, N., Bauer, M., Durant, J., Allansdottir, A., Bonfadelli, H., Boy, D., de Cheveigné, S., Fjaestad, B., Gutteling, J. M., Hampel, J., Jelsøe, E., Jesuino, J. C., Kohring, M., Kronberger, N., Midden, C., Nielsen, T. H., Przystalski, A., Rusanen, T., ... Wagner, W. (2000). Biotechnology and the European public. *Nature Biotechnology*, 18(9), 935–938. <https://doi.org/10.1038/79403>
- Hong, J., Shi, Q., Biswas, S., Jiang, S.-C., & Shi, J. (2021). Moving genome edited crops forward from the laboratory bench to the kitchen table. *Food Control*, 122, 107790. <https://doi.org/10.1016/j.foodcont.2020.107790>

- Hunt, S., & Frewer, L. J. (2001). Trust in sources of information about genetically modified food risks in the UK. *British Food Journal*, 103(1), 46–62.
<https://doi.org/10.1108/00070700110383019>
- ISAAA. (2020). *Brief 55: Global Status of Commercialized Biotech/GM Crops: 2019*. International Service for the Acquisition of Agri-biotech Applications.
- Ishwaran, H., & Kogalur, U. B. (2019). *Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC)*. CRAN. <http://www.est.colpos.mx/R-mirror/web/packages/randomForestSRC/randomForestSRC.pdf>
- Ishwaran, H., Kogalur, U. B., Gorodeski, E. Z., Minn, A. J., & Lauer, M. S. (2010). High-Dimensional Variable Selection for Survival Data. *Journal of the American Statistical Association*, 105(489), 205–217. <https://doi.org/10.1198/jasa.2009.tm08622>
- Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, 64(5), 402–402. <https://doi.org/10.4097/kjae.2013.64.5.402>
- Karpiévitch, Y. V., Hill, E. G., Leclerc, A. P., Dabney, A. R., & Almeida, J. S. (2009). An Introspective Comparison of Random Forest-Based Classifiers for the Analysis of Cluster-Related Data by Way of RF++. *PLoS ONE*, 4(9), e7087–e7087.
<https://doi.org/10.1371/journal.pone.0007087>
- Kato-Nitta, N., Maeda, T., Inagaki, Y., & Tachikawa, M. (2019). Expert and public perceptions of gene-edited crops: Attitude changes in relation to scientific knowledge. *Palgrave Communications*, 5(1), 137. <https://doi.org/10.1057/s41599-019-0328-4>
- Kennedy, B., & Thigpen, C. L. (2020). *Many publics around world doubt safety of genetically modified foods*. Pew Research Center. <https://www.pewresearch.org/fact-tank/2020/11/11/many-publics-around-world-doubt-safety-of-genetically-modified-foods/>
- Knijnenburg, T. A., Wessels, L. F. A., Reinders, M. J. T., & Shmulevich, I. (2009). Fewer permutations, more accurate P-values. *Bioinformatics*, 25(12), i161–i168.
<https://doi.org/10.1093/bioinformatics/btp211>
- Kogalur, U. B., & Ishwaran, H. (2020). *Random Forests for Survival, Regression, and Classification—Theory and Specifications*.
<https://kogalur.github.io/randomForestSRC/theory.html>
- Ledford, H. (2019). Gene-edited animal creators look beyond US market. *Nature*, 566(7745), 433–434. <https://doi.org/10.1038/d41586-019-00600-4>
- Legge Jr., J. S., & Durant, R. F. (2010). Public Opinion, Risk Assessment, and Biotechnology: Lessons from Attitudes toward Genetically Modified Foods in the European Union. *Review of Policy Research*, 27(1), 59–76. <https://doi.org/10.1111/j.1541-1338.2009.00427.x>
- Lloyd's Register Foundation & Gallup. (2020). *The Lloyd's Register Foundation World Risk Poll*. Lloyd's Register Foundation.

- Lloyds Register Foundation & Gallup. (2020). *The Llyod's Register Foundation World Risk Poll—2019 Methodology*.
https://wrp.lrfoundation.org.uk/LRF_WorldRiskReport_MethodologyAppendix.pdf
- Ma, X., Mau, M., & Sharbel, T. F. (2018). Genome Editing for Global Food Security. *Trends in Biotechnology*, 36(2), 123–127. <https://doi.org/10.1016/j.tibtech.2017.08.004>
- Malyska, A., Bolla, R., & Twardowski, T. (2016). The Role of Public Opinion in Shaping Trajectories of Agricultural Biotechnology. *Trends in Biotechnology*, 34(7), 530–534. <https://doi.org/10.1016/j.tibtech.2016.03.005>
- Marris, C. (2001). Public views on GMOs: Deconstructing the myths: Stakeholders in the GMO debate often describe public opinion as irrational. But do they really understand the public? *EMBO Reports*, 2(7), 545–548. <https://doi.org/10.1093/embo-reports/kve142>
- Martin, D. P. (2015). *Efficiently Exploring Multilevel Data with Recursive Partitioning*. <https://dpmartin42.github.io/extras/dissertation.pdf>
- Martin, H. M., Durr, D., Smith, L. M., Finke, R., & Cherry, A. (2017). Analysis of GMO food products companies: Financial risks and opportunities in the global agriculture industry. *African J. of Economic and Sustainable Development*, 6(1), 1. <https://doi.org/10.1504/AJESD.2017.082813>
- Mayer, M. (2019). *Package 'missRanger.'* <https://cran.rstudio.com/web/packages/missRanger/missRanger.pdf>
- McFarlane, G. R., Salvesen, H. A., Sternberg, A., & Lillico, S. G. (2019). On-Farm Livestock Genome Editing Using Cutting Edge Reproductive Technologies. *Frontiers in Sustainable Food Systems*, 3, 106. <https://doi.org/10.3389/fsufs.2019.00106>
- McMillan, G. S. (2009). Gender differences in patenting activity: An examination of the US biotechnology industry. *Scientometrics*, 80(3), 683–691. <https://doi.org/10.1007/s11192-008-2101-0>
- Menz, J., Modrzejewski, D., Hartung, F., Wilhelm, R., & Sprink, T. (2020). Genome Edited Crops Touch the Market: A View on the Global Development and Regulatory Environment. *Frontiers in Plant Science*, 11, 586027. <https://doi.org/10.3389/fpls.2020.586027>
- Molina, M., & Garip, F. (2019). Machine Learning for Sociology. *Annual Review of Sociology*, 45(1), 27–45. <https://doi.org/10.1146/annurev-soc-073117-041106>
- Myers, S. S., Smith, M. R., Guth, S., Golden, C. D., Vaitla, B., Mueller, N. D., Dangour, A. D., & Huybers, P. (2017). Climate Change and Global Food Systems: Potential Impacts on Food Security and Undernutrition. *Annual Review of Public Health*, 38(1), 259–277. <https://doi.org/10.1146/annurev-publhealth-031816-044356>
- Ojala, M., & Garriga, G. C. (2010). Permutation Tests for Studying Classifier Performance. *Journal of Machine Learning Research*, 11, 1833–1863.
- Paarlberg, R. (2010). GMO foods and crops: Africa's choice. *New Biotechnology*, 27(5), 609–613. <https://doi.org/10.1016/j.nbt.2010.07.005>

- Paul, M. J., Nuccio, M. L., & Basu, S. S. (2018). Are GM Crops for Yield and Resilience Possible? *Trends in Plant Science*, 23(1), 10–16.
<https://doi.org/10.1016/j.tplants.2017.09.007>
- Shew, A. M., Nalley, L. L., Snell, H. A., Nayga, R. M., & Dixon, B. L. (2018). CRISPR versus GMOs: Public acceptance and valuation. *Global Food Security*, 19, 71–80.
<https://doi.org/10.1016/j.gfs.2018.10.005>
- Springmann, M., Mason-D'Croz, D., Robinson, S., Garnett, T., Godfray, H. C. J., Gollin, D., Rayner, M., Ballon, P., & Scarborough, P. (2016). Global and regional health effects of future food production under climate change: A modelling study. *The Lancet*, 387(10031), 1937–1946. [https://doi.org/10.1016/S0140-6736\(15\)01156-3](https://doi.org/10.1016/S0140-6736(15)01156-3)
- Steinwand, M. A., & Ronald, P. C. (2020). Crop biotechnology and the future of food. *Nature Food*, 1(5), 273–283. <https://doi.org/10.1038/s43016-020-0072-3>
- Stekhoven, D. J., & Buhlmann, P. (2012). MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112–118.
<https://doi.org/10.1093/bioinformatics/btr597>
- Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(1), 25–25. <https://doi.org/10.1186/1471-2105-8-25>
- Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14(4), 323–348.
<https://doi.org/10.1037/a0016973>
- Strobl, C., & Zeileis, A. (2008). *Danger: High Power! – Exploring the Statistical Properties of a Test for Random Forest Variable Importance*.
<https://core.ac.uk/download/pdf/12163053.pdf>
- Tait-Burkard, C., Doeschl-Wilson, A., McGrew, M. J., Archibald, A. L., Sang, H. M., Houston, R. D., Whitelaw, C. B., & Watson, M. (2018). Livestock 2.0 – genome editing for fitter, healthier, and more productive farmed animals. *Genome Biology*, 19(1), 204.
<https://doi.org/10.1186/s13059-018-1583-1>
- Tang, F., & Ishwaran, H. (2017). Random forest missing data algorithms. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 10(6), 363–377.
<https://doi.org/10.1002/sam.11348>
- U.S. Department of Agriculture. (2021). *Agricultural Biotechnology Glossary*.
<https://www.usda.gov/topics/biotechnology/biotechnology-glossary>
- Weisberg, S. M., Badgio, D., & Chatterjee, A. (2017). A CRISPR New World: Attitudes in the Public toward Innovations in Human Genetic Modification. *Frontiers in Public Health*, 5, 117. <https://doi.org/10.3389/fpubh.2017.00117>
- Woo, J. W., Kim, J., Kwon, S. I., Corvalán, C., Cho, S. W., Kim, H., Kim, S.-G., Kim, S.-T., Choe, S., & Kim, J.-S. (2015). DNA-free genome editing in plants with preassembled

- CRISPR-Cas9 ribonucleoproteins. *Nature Biotechnology*, 33(11), 1162–1164.
<https://doi.org/10.1038/nbt.3389>
- World Bank. (2021a). *Agriculture, forestry, and fishing, value added (% of GDP)*.
<https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS>
- World Bank. (2021b). *GDP per capita (current US\$)*.
<https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD>
- World Bank. (2021c). *Scientific and technical journal articles*. World Bank Development Indicators. <https://data.worldbank.org/indicator/IP.JRN.ARTC.SC>
- World Bank. (2021d). *The Worldwide Governance Indicators (WGI) project*.
<https://info.worldbank.org/governance/wgi/>
- World Food Programme. (2020). *The power of gender equality for food security*.
<https://docs.wfp.org/api/documents/WFP-0000112500/download/>
- Zhang, Y., Pribil, M., Palmgren, M., & Gao, C. (2020). A CRISPR way for accelerating improvement of food crops. *Nature Food*, 1(4), 200–205. <https://doi.org/10.1038/s43016-020-0051-8>