Deforestation reduces fruit and vegetable consumption in rural Tanzania

Charlotte M. Halla, Laura Vang Rasmussen, Bronwen Powel, Cecilie Dyngelanda, Suhyun Jung, and Rasmus Skov Olesen

"Department of Geosciences and Natural Resource Management, University of Copenhagen, 1350 Copenhagen, Denmark; Department of Geography, The Pennsylvania State University, State College, PA 16801; Department of Agricultural Sciences, Inland Norway University of Applied Sciences, 2318 Hamar, Norway; and Division of Resource Economics and Management, West Virginia University, Morgantown, WV 26506

Edited by Patrick Meyfroidt, Earth and Life Institute, Georges Lemaître Earth and Climate Center, Universite Catholique de Louvain, Louvain-La-Neuve, Belgium; received June 30, 2021; accepted January 18, 2022 by Editorial Board Member Carl Folke

Strategies to improve food and nutrition security continue to promote increasing food via agricultural intensification. Little (if any) consideration is given to the role of natural landscapes such as forests in meeting nutrition goals, despite a growing body of literature that shows that having access to these landscapes can improve people’s diets, particularly in rural areas of low- and middle-income countries. In this study, we tested whether deforestation over a 5-year period (2008–2013) affected people’s dietary quality in rural Tanzania using a modeling approach that combined two-way fixed-effects regression analysis with covariate balancing generalized propensity score (CBGPS) weighting which allowed for causal inferences to be made. We found that, over the 5 y, deforestation caused a reduction in household fruit and vegetable consumption and thus vitamin A adequacy of diets. The average household member experienced a reduction in fruit and vegetable consumption of 14 g d⁻¹, which represented a substantial proportion (11%) of average daily intake. Conversely, we found that forest fragmentation over the survey period led to an increase in consumption of these foods and dietary vitamin A adequacy. This study finds a causal link between deforestation and people’s dietary quality, and the results have important implications for policy makers given that forests are largely overlooked in strategies to improve nutrition, but offer potential “win–wins” in terms of meeting nutrition goals as well as conservation and environmental goals.

deforestation | diet quality | wild foods

The challenge of achieving food and nutrition security for the world’s growing population while also minimizing and reversing damage to the natural environment is unprecedented. The dominant narrative on how to achieve food and nutrition security continues to be centered on intensifying agricultural production to produce more food (1–3). While agricultural intensification is undoubtedly a key reason we have kept pace with food demands and ended hunger for millions of people over the past decades, it has led to a preoccupation with dietary energy (calories), and thus the production of staple grains which provide the majority of calories globally (4, 5). The focus on staple foods has resulted in dietary quality and diversity being overlooked, despite the fact that far more people suffer from micronutrient deficiency than undernourishment (6–8). Likewise, agricultural intensification is a leading driver of environmental degradation (9–11). There has been much research in recent years examining the impact of different diets on land use (12–15), but less attention has been given to the reverse of this relationship: How do landscapes affect diets? A growing body of literature has examined this relationship with a focus on the linkages between forests and diets in low- and middle-income countries. This relatively new field of research has important implications for strategies to achieve food and nutrition security worldwide, particularly for rural areas in low- and middle-income countries where there are strong connections between livelihoods and landscapes, and undernourishment is most prevalent.

Forests provide critical ecosystem services that benefit human populations in various ways, such as the provision of food and fiber, and climate and water regulation (16), with an estimated 1.5 billion forest-proximate people worldwide (i.e., living within 5 km of a forest) (17). Forests can improve people’s diets via four key pathways (18, 19). The most direct way is via the provision of wild forest foods, which most often include fruits, vegetables, mushrooms, and animal products (i.e., bushmeat and insects), all of which tend to be high in essential micronutrients (20–22). The second pathway is via income generation from the sale of forest foods and other non-forest timber forest products (NTFPs), which can improve livelihoods and facilitate the purchase of nutritious foods from markets (23, 24). The third pathway is via the flow of ecosystem services from forests into surrounding agricultural landscapes (e.g., forests can contribute to soil formation and nutrient cycling, and increase pollination) which can increase and/or diversify production (25). The final pathway is the provision of fuelwood for cooking, which is a key (but often overlooked) pathway that can improve nutrition by facilitating the preparation of a range of foods, particularly those with long cooking times (26, 27).

Significance

Two billion people across the planet suffer from nutrient deficiencies. Dietary diversification is key to solving this problem, yet many food and nutrition security policies, especially in low- and middle-income countries, still focus on increasing agricultural production and access to sufficient calories as the main solution. But calories are not all equal. Here, we show how deforestation in Tanzania caused a reduction in fruit and vegetable consumption (of 14 g per person per day) and thus vitamin A adequacy of diets. Using a combination of regression and weighting analyses to generate quasi-experimental quantitative estimates of the impacts of deforestation on people’s food intake, our study establishes a causal link between deforestation and people’s dietary quality.

Author contributions: C.M.H. and L.V.R. designed research; C.M.H. and L.V.R. performed research; C.M.H., L.V.R., B.P., C.D., and S.J. contributed new analytic tools; C.M.H., L.V.R., and C.D. analyzed data; R.O.S. contributed to the background literature review; and C.M.H. wrote the paper with contributions from all authors.

The authors declare no competing interest.

This article is a PNAS Direct Submission. P.M. is a guest editor invited by the Editorial Board.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1 To whom correspondence may be addressed. Email: Chat@lign.ku.dk.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2112063119/-/DCSupplemental.

Published March 1, 2022.
The majority of studies have found a positive relationship between living near (having access to) forests and several measures of diet, nutrition, and food security outcomes. Most studies use metrics of diet quality such as dietary diversity scores or consumption of certain nutritious food groups. Very few studies have examined more detailed measures of dietary quality such as energy and nutrient intakes (28–30), and only one study has examined these in relation to forest cover using multivariate regression (31). Moreover, the majority of studies examine the relationship between forests and diet quality at a single point in time. Two studies have examined the relationship between diets and previous forest loss (32, 33), but no studies, to date, have used longitudinal data to understand concurrent changes in forests and diets over time. In this sense, most studies have only been able to identify associations between forests and diets as opposed to causal relationships. Furthermore, only one study, to date, has examined how the spatial arrangement of forests (as opposed to just forest amount) can affect people’s diets (34), finding that forest configuration may be as important as forest amount for dietary quality.

This study aimed to advance the current knowledge on the forest–diet relationship in three main ways:

1) By using panel data and a rigorous estimation method which combines covariate balancing generalized propensity score (CBGPS) weighting with two-way fixed-effects regression, we were able to test the causal impact of forest changes on diets, which no studies, to our knowledge, have done. We were also able to explore the causal mechanisms by which forest cover change is hypothesized to affect people’s diets (the direct consumption pathway, the income pathway, and the ecosystem services pathway).

2) Most existing studies rely on measures such as dietary diversity scores and consumption of nutritious food groups as proxies for overall diet quality. In addition to these, we also quantified household energy and nutrient adequacy levels in order to gain a better understanding of how forests can affect people’s diets.

3) We considered not just forest amount but also the spatial arrangement of forests in relation to diet quality, which only one study has done, to date (34). Thus, this study aimed to extend this research to examine whether changes in forest configuration [in terms of fragmentation (35)] were related to people’s dietary quality.

Results

Descriptive Statistics. Summary statistics for the key outcome and explanatory variables included in our modeling efforts are shown in Table 1 (also see SI Appendix, Table S1, which shows medians for all key variables). Mean forest cover (hectares) across the clusters (note that a cluster is a sampling unit used by the Living Standards Measurement Study [LSMS] normally corresponding to a village) decreased over time with an average loss of 171 ha over the 5-y period (a statistically significant reduction). The majority of clusters (94%) saw a decrease in forest cover, but a few had no net change (2%) or a small gain in cover (4%) (Fig. 1). Forest fragmentation (measured as the number of individual forest patches within a cluster) increased over time (statistically significant), with an average increase of 27 patches but with significant variation across clusters. An increase in patches was apparent in 44% of clusters, while 6% had no change and 50% exhibited a decrease in the number of patches (Fig. 1).

There was a general decline in dietary quality over the 5-y period. Dietary diversity increased marginally between waves one and two, and then decreased between waves two and three (but the overall change between waves one and three was not statistically significant). Mean fruit and vegetable consumption was very low in all waves (~130 g per person per day, with less than 5% of households meeting recommended intakes of 400 g per day), and there was a marginal decline over the 5-y period (not statistically significant). There were more substantial declines in energy and nutrient adequacy levels over the survey period which were all statistically significant. Average energy, protein, and micronutrient adequacy levels decreased by around 10% in each case between waves one and three, and thus the percentage of households meeting recommended intakes showed a similar pattern.

Table 1. Summary statistics for the key dependent and independent variables in each wave of the panel data

| Variable                          | 2008–2009 Mean (SD) | 2010–2011 Mean (SD) | 2012–2013 Mean (SD) |
|----------------------------------|---------------------|---------------------|---------------------|
| Forest cover (ha)                | 8,189.8 (7,687.5)   | 8,189.5 (7,690.5)   | 8,018.5 (7,601.7)***|
| Number of forest patches         | 1,158.9 (932.9)     | 1,159.1 (933.3)     | 1,183 (960.4)***    |
| MDDS                             | 5.7 (1.5)           | 5.8 (1.5)           | 5.7 (1.5)           |
| Fruit and vegetable consumption (g/AME/d) | 132.4 (141.7)     | 129.5 (105.4)       | 131.3 (118.7)       |
| Energy intake (kcalis/AME/d)     | 2,824.9 (1110.5)    | 2,367.9 (961.8)     | 2,293.2 (996.9)***  |
| Protein intake (g/AME/d)         | 74.4 (36.8)         | 62.4 (30.3)         | 60.9 (32.3)***      |
| Iron intake (mg/AME/d)           | 21.2 (10.8)         | 16.9 (8.5)          | 16.6 (8.8)***       |
| Zinc intake (mg/AME/d)           | 11.9 (5.9)          | 9.9 (4.8)           | 9.8 (5.2)***        |
| Vitamin A intake (RAE µg/AME/d)  | 1,289.4 (1780.3)    | 966.9 (1278.3)      | 1,079.2 (1456)***   |
| Energy adequacy ratio (%)        | 85.4 (19.6)         | 78.4 (21.2)         | 75.4 (22.4)***      |
| Protein adequacy ratio (%)       | 93.9 (14.3)         | 91.4 (16.3)         | 88.7 (18.8)***      |
| Iron adequacy ratio (%)          | 53.3 (26.3)         | 43.9 (22.4)         | 43.4 (23.4)***      |
| Zinc adequacy ratio (%)          | 71.1 (26.1)         | 62.6 (24.2)         | 61.1 (25.8)***      |
| Vitamin A adequacy ratio (%)     | 75.2 (30.5)         | 74.8 (28.9)         | 74.4 (29.9)***      |

Households meeting fruit and vegetable recommendations (%) | 4.8 | 2.6 | 3.7 |
Households meeting energy requirements (%) | 49.3 | 31.1 | 27.3 |
Households meeting protein requirements (%) | 76.9 | 67.8 | 61.1 |
Households meeting iron requirements (%) | 8.9 | 3.3 | 3.6 |
Households meeting zinc requirements (%) | 27.2 | 13.8 | 14.6 |
Households meeting vitamin A requirements (%) | 48.3 | 44.3 | 45.5 |

Values in the upper part of the table are means with SDs in parentheses, while values in the lower part are proportions (percent). Asterisks denote whether the changes between waves one and three were statistically significant (**P < 0.01).
Deforestation reduces fruit and vegetable consumption in rural Tanzania

Hall et al.

PNAS | 3 of 9

http://www.pnas.org

Deforestation reduced fruit and vegetable consumption and vitamin A adequacy. Our results have established a causal relationship between forest cover loss (i.e., deforestation) and a reduction in people's fruit and vegetable consumption. The specific fruit and vegetable categories responsible for this decline were "spinach, cabbage, and other green vegetables" and "mangoes, avocados, and other fruits" (22), suggesting that these food groups were driving the reduction in fruit and vegetable consumption. Further investigation showed that the increase in fruit and vegetable consumption was being driven by the categories "onions, tomatoes, carrots, green pepper, and other viungo" and "spinach, cabbage, and other green vegetables," but a negative association was found between forest patches and "citrus fruits" (SI Appendix, Table S3).

In terms of the control variables, household size was positively associated with dietary diversity but negatively associated with fruit and vegetable consumption and all dietary adequacy variables, suggesting that larger households have more-diverse diets but less adequate intakes overall. Age of the household head was negatively associated with dietary diversity and fruit and vegetable consumption, suggesting that households with older heads may have generally poorer-quality diets than households headed by younger members. Households headed by women had significantly higher dietary diversity and iron adequacy. Having a high wealth level was positively associated with energy and protein adequacy, and having a middle wealth level was positively associated with all dietary adequacy variables except for vitamin A. Interestingly, dietary diversity and fruit and vegetable consumption were not affected by wealth. Having a household head educated at primary level was associated with significantly higher energy adequacy than households where the head had no education.

Crop count was positively associated with dietary diversity and protein adequacy but was negatively associated with vitamin A adequacy. No relationship was found between crop count and fruit and vegetable consumption. Livestock ownership was positively associated with energy, zinc, and vitamin A adequacy but, surprisingly, not iron adequacy. This suggests that households that owned livestock consumed more animal products, given these foods tend to be energy dense and high in these nutrients (but we acknowledge other pathways by which livestock ownership can contribute to dietary intake, such as via income generation). Lastly, seasonality was only a significant determinant of protein adequacy, suggesting that more protein-rich foods were eaten during the rainy season.

Discussion

Deforestation Reduced Fruit and Vegetable Consumption and Vitamin A Adequacy. Our results have established a causal relationship between forest cover loss (i.e., deforestation) and a reduction in people's fruit and vegetable consumption. The specific fruit and vegetable categories responsible for this decline were "spinach, cabbage, and other green vegetables" and "mangoes, avocados, and other fruits" (22), suggesting that these food groups were driving the reduction in fruit and vegetable consumption. Further investigation showed that the increase in fruit and vegetable consumption was being driven by the categories "onions, tomatoes, carrots, green pepper, and other viungo" and "spinach, cabbage, and other green vegetables," but a negative association was found between forest patches and "citrus fruits" (SI Appendix, Table S3).

In terms of the control variables, household size was positively associated with dietary diversity but negatively associated with fruit and vegetable consumption and all dietary adequacy variables, suggesting that larger households have more-diverse diets but less adequate intakes overall. Age of the household head was negatively associated with dietary diversity and fruit and vegetable consumption, suggesting that households with older heads may have generally poorer-quality diets than households headed by younger members. Households headed by women had significantly higher dietary diversity and iron adequacy. Having a high wealth level was positively associated with energy and protein adequacy, and having a middle wealth level was positively associated with all dietary adequacy variables except for vitamin A. Interestingly, dietary diversity and fruit and vegetable consumption were not affected by wealth. Having a household head educated at primary level was associated with significantly higher energy adequacy than households where the head had no education.

Crop count was positively associated with dietary diversity and protein adequacy but was negatively associated with vitamin A adequacy. No relationship was found between crop count and fruit and vegetable consumption. Livestock ownership was positively associated with energy, zinc, and vitamin A adequacy but, surprisingly, not iron adequacy. This suggests that households that owned livestock consumed more animal products, given these foods tend to be energy dense and high in these nutrients (but we acknowledge other pathways by which livestock ownership can contribute to dietary intake, such as via income generation). Lastly, seasonality was only a significant determinant of protein adequacy, suggesting that more protein-rich foods were eaten during the rainy season.

Discussion

Deforestation Reduced Fruit and Vegetable Consumption and Vitamin A Adequacy. Our results have established a causal relationship between forest cover loss (i.e., deforestation) and a reduction in people's fruit and vegetable consumption. The specific fruit and vegetable categories responsible for this decline were "spinach, cabbage, and other green vegetables" and "mangoes, avocados, and other fruits" (22), suggesting that these food groups were driving the reduction in fruit and vegetable consumption. Further investigation showed that the increase in fruit and vegetable consumption was being driven by the categories "onions, tomatoes, carrots, green pepper, and other viungo" and "spinach, cabbage, and other green vegetables," but a negative association was found between forest patches and "citrus fruits" (SI Appendix, Table S3).

In terms of the control variables, household size was positively associated with dietary diversity but negatively associated with fruit and vegetable consumption and all dietary adequacy variables, suggesting that larger households have more-diverse diets but less adequate intakes overall. Age of the household head was negatively associated with dietary diversity and fruit and vegetable consumption, suggesting that households with older heads may have generally poorer-quality diets than households headed by younger members. Households headed by women had significantly higher dietary diversity and iron adequacy. Having a high wealth level was positively associated with energy and protein adequacy, and having a middle wealth level was positively associated with all dietary adequacy variables except for vitamin A. Interestingly, dietary diversity and fruit and vegetable consumption were not affected by wealth. Having a household head educated at primary level was associated with significantly higher energy adequacy than households where the head had no education.

Crop count was positively associated with dietary diversity and protein adequacy but was negatively associated with vitamin A adequacy. No relationship was found between crop count and fruit and vegetable consumption. Livestock ownership was positively associated with energy, zinc, and vitamin A adequacy but, surprisingly, not iron adequacy. This suggests that households that owned livestock consumed more animal products, given these foods tend to be energy dense and high in these nutrients (but we acknowledge other pathways by which livestock ownership can contribute to dietary intake, such as via income generation). Lastly, seasonality was only a significant determinant of protein adequacy, suggesting that more protein-rich foods were eaten during the rainy season.
consumption was just 130 g per capita per day across the panel waves, and such a small proportion of household members consumed the recommended amounts of 400 g per capita per day (<5% in each wave).

In terms of vitamin A, further analyses (SI Appendix, section B) showed that the reduction in household vitamin A adequacy was a result of the decrease in fruit and vegetable consumption and was not being driven by a reduction in any other vitamin A-rich foods (such as sweet potato or palm oil). Similarly, the specific fruit and vegetable categories that were affected by forest cover have high vitamin A contents (SI Appendix, Table S4). For example, the category “mangoes, avocados, and other fruits” has 12 times higher vitamin A content than “ripe bananas,” and 4 times higher vitamin A content than “citrus fruits.” Our findings relating to fruit and vegetable consumption and vitamin A further advance the two existing studies that have found previous forest loss to be associated with lower consumption of nutritious foods (32, 33). Specifically, our findings add to Johnson et al. (32), who found that children in Malawi who experienced forest loss over a 10-y period were 29% less likely to consume vitamin A–rich foods than children who did not experience a net loss of forest cover. Our findings on vitamin A are also in line with some studies that found living near to, and having access to, forest landscapes was beneficial for vitamin A intake (28, 31, 37).

While we can ascertain that, in our study, deforestation caused a reduction in consumption of certain fruits and vegetables (and thus vitamin A adequacy), we cannot determine the exact causal mechanisms given the nature of our data. Yet, given the types of fruits and vegetables affected by forest cover in this study, and that other studies have found the direct provision of wild fruits and vegetables to be a key forest–diet pathway in Tanzania, it is likely that deforestation reduced the availability of these foods for direct collection and consumption. For example, Powell et al. (22, 28) found that women and children in the East Usambara mountains in Tanzania sourced a wide range of nutrient-dense, wild plant and animal foods from surrounding landscapes (including forests), which made important contributions to their vitamin A, vitamin C, and iron intakes. Similarly, Msuya et al. (38) found that communities in Uluguru North and the West Usambara Mountains in Tanzania reported consumption of 114 indigenous forest food plant species. In addition, an earlier study (39) in the Lushoto District of northeastern Tanzania found that households frequently consumed green leafy vegetables, a high proportion of which came from the forest.

We also carried out investigation into other potential pathways (the income pathway and ecosystem services pathway) but did not find convincing results (SI Appendix, section C). For example, deforestation could have reduced people’s ability to collect and sell NTFPs, ultimately affecting their income and ability to purchase fruits and vegetables at markets (23, 40, 41). However, we found no relationship between forest cover change and household expenditures, nor did we find any relationship between household expenditures and the consumption of those fruit and vegetable categories that were related to forest cover change. Additionally, agricultural production in areas surrounding forests could have been affected by deforestation via a loss of ecosystem services such as pollination, ultimately leading to a reduction in fruit and vegetable production and consumption (25, 42–45). However, we did not find significant associations between household plot size (i.e., the total land area cultivated by households in each wave) and the consumption of those fruit and vegetable categories related to forest cover. Thus, we did not find evidence to suggest that the income or ecosystem services pathways led to the reduction in fruit and vegetable consumption in our study, which

Table 2. Results from the two-way fixed-effects regression models including the CBGPS weights

|          | MDDS | Fruit and vegetable consumption | Energy adequacy | Protein adequacy | Iron adequacy | Zinc adequacy | Vitamin A adequacy |
|----------|------|---------------------------------|----------------|-----------------|--------------|---------------|--------------------|
| Forest cover | NS   | 0.08 (0.02)***                  | NS             | NS              | NS           | NS            | 0.01 (0.005)**     |
| Forest patches | NS   | 0.12 (0.03)***                  | NS             | NS              | NS           | NS            | 0.02 (0.007)**     |
| Household size | 0.07 (0.02)**                  | -9.22 (1.87)*** | -1.24 (0.32)*** | -1.01 (0.23)*** | -2.79 (0.31)*** | -2.40 (0.34)*** | -1.25 (0.40)**     |
| Age     | -0.02 (0.01)**                  | -1.97 (0.68)**   | NS             | NS              | NS           | NS            | NS                 |
| Sex (female) | 0.46 (0.19)**                  | NS              | NS             | NS              | NS           | NS            | NS                 |
| Wealth (low) | NS   | 5.49 (1.01)***                  | 1.91 (0.74)*   | NS              | NS           | NS            | NS                 |
| Wealth (middle) | NS   | 4.73 (0.88)***                  | 3.30 (0.64)*** | 2.09 (0.85)*   | 3.47 (0.94)*** | NS            | NS                 |
| Education (primary) | NS   | NS                            | NS             | NS              | NS           | NS            | NS                 |
| Education (secondary) | NS   | NS                           | NS             | NS              | NS           | NS            | NS                 |
| Crop count | 0.06 (0.02)*                   | NS              | NS             | NS              | NS           | NS            | NS                 |
| Livestock ownership | NS   | 3.45 (1.16)**                  | NS             | NS              | NS           | NS            | NS                 |
| Season (rainy) | 2.38 (1.17)*                   | NS              | NS             | NS              | NS           | NS            | NS                 |

Values are model coefficients with test statistics in parentheses. *NS* denotes not significant. *<0.05; **<0.01; ***<0.001.

Fig. 2. Coefficient plots summarizing the regression outputs for models run between forest cover change and consumption of each fruit and vegetable category (grams per AME per day) over the study period (2008–2013). ***<0.001.
Deforestation reduces fruit and vegetable consumption in rural Tanzania

Deforestation Does Not Affect Dietary Diversity Scores. The absence of a relationship between forest cover and dietary diversity in this study was inconsistent with most other studies that have found a positive association between the amount of forest cover in people’s surroundings and their dietary diversity (e.g., refs. 32–34 and 46–48). A possible explanation for this is the pathways by which forests contribute to diets in Tanzania. Studies have shown that there are different pathways in different places (as discussed previously and as summarized by Baum et al. (48)). Given that crop count was positively associated with dietary diversity and protein adequacy, and that crop count is also a proxy for income generated from their sale to purchase a more diverse range of foods, as well as protein-rich foods (such as fish and other animal products) (23, 40, 41).

Forest Fragmentation May Improve Dietary Quality. Our findings related to forest patches and people’s diets offer new insights, with only one study, to our knowledge, having empirically examined associations between forest cover (as opposed to just forest amount) and dietary quality (34). The study by Rasmussen et al. (34) found that people living in landscapes with more forest patches were significantly more likely to consume fruits than people living in less fragmented landscapes. Specifically in Tanzania, the mean predicted probability of consuming fruits increased by a factor of 1.5 from the first to the second quintile of forest patches. Our results are in line with these findings, as we found that households who experienced an increase in fragmentation had an increased intake of overall fruit and vegetable consumption (driven by an increase in the categories “onions, tomatoes, carrots, green pepper, and other vegetables”), as well as more adequate vitamin A intakes. Yet, unlike for forest cover loss, these findings were not causal, and the effect sizes were much smaller.

Given the lack of studies on how forest configuration may affect people’s diets, it is hard to ascertain why an increase in forest patches was associated with an increase in fruit and vegetable consumption. However, there are a number of possible explanations as outlined by Rasmussen et al. (34), and Friant et al. (49): 1) Households are more likely to collect wild forest foods from smaller blocks of forest due to better access (23, 2) many forest foods actually come from forest edges, and more fragmented forests have greater edge length which could improve access to wild foods (20, 22, 26), 3) smaller patches of forest may actually be “managed” (we were not able to discern forest type using the Hansen dataset), and are thus maintained to produce certain foods (50), 4) less fragmented forests may have restricted access for conservation reasons (3), and 5) smaller blocks of forest may lead to more effective pollination of nearby domestic food crops, leading to an increased consumption of these foods (51).

Study Limitations. There are important limitations that should be considered when interpreting the results of this study. Firstly, there are limitations to using the LSMS food consumption data to estimate nutrient intake and adequacy, as these metrics are usually assessed from an individual 24-h dietary recall with built-in methods to improve recall and portion size estimations and reduce omissions and bias. Thus, using the LSMS data could have resulted in overestimation or underestimation of food consumed by households. Overestimation could have occurred due to some households “bulk buying” certain produce (i.e., bags of grain) but only eating a small amount of it during the recall period. However, most of these cases would likely have been removed when we adjusted for “implausible” calorie intakes. Similarly, while we adjusted the consumed food weights for inedible portions, the LSMS does not account for food waste at the household level. Yet, this is unlikely to be significant, as household-level food waste in countries such as Tanzania is estimated to be very low (52). Underestimation of food consumption could also have occurred due to the underreporting of foods consumed outside the household, which the LSMS does not account for. Given that the food consumption questionnaire was completed by only one household member, it is unlikely that all foods consumed by all members of the household were accounted for. Foods sourced from the forest by certain household members might thus go unrecorded. However, even 24-h recall surveys struggle with this issue; for example, Fleuret (39) found major underreporting of fruits, especially when consumed outside of the household. Likewise, given that the food consumption survey was conducted using a predetermined list of food and drink items, it is possible that some foods consumed by the household were not included in the list and were thus not reported.

Potential inaccuracies in the estimation of nutrient intake were unavoidable given the nature of the LSMS data. In some cases, the food consumption data did not include individual food items but rather groups of similar foods (e.g., citrus fruits). Determining the nutrient composition of these required taking averages (an average was taken of lemons, limes, oranges, tangerines, and grapefruit, in the case of “citrus fruit”). Calculating nutrient intake, whether from household- or individual-level data, is also limited by a dearth of food composition information on nutrient content (53).

Despite the limitations of using household consumption and expenditure surveys (HCES) such as the LSMS to estimate dietary intake, some studies have promoted these data sources for this purpose given the scarcity of finer-scale food consumption data at national levels (54–56). For example, Bermudez et al. (54), used HCES data from Bangladesh to estimate apparent
intakes of calories, vitamin A, iron, and zinc across 10,080 households using adult male equivalent (AME) values (as was done in this study). The results were remarkably close to estimates from the World Food Program and other reported research for the same period. Thus, while we acknowledge the limitations of the LSMS data and the benefits of more-direct methods, the results are valuable, particularly given the national coverage which could guide the identification of hotspots of vulnerable households to target nutritional interventions. Importantly, we believe the methods used have internal reliability, allowing us to compare across households and time points, even if there are any of the above forms of systematic bias in the methods.

Conclusions. Our findings have policy relevance in terms of future strategies for improving and protecting food and nutrition security, particularly in rural areas of low- and middle-income countries. Our findings support the growing body of literature that links biodiverse landscapes such as those that include forests with better nutritional and overall health outcomes for local communities. Yet, national strategies to improve food security are often still focused on agricultural intensification and increasing yields of staple crops, with little or no attention given to the role of forests or wild foods. While increased agricultural production will inevitably play an important role in meeting the food needs of a growing population, the focus on staple crop yields does little to address issues around dietary quality. For example, insufficient fruit and vegetable intake is considered to be a leading risk factor for chronic disease globally, and a major barrier in achieving healthy diets (57). The importance of forests for fruit and vegetable consumption may be even greater in arid regions where the cultivation of additional fruits and vegetables may be limited due to water access (58).

Given the growing body of literature that links forests with fruit and vegetable consumption and overall improved dietary quality, it is worrying that national (and international) food and agriculture policies rarely attend to the role of forests in helping to improve nutrition. For example, the most recent report from the High Level Panel of Experts on Food Security and Nutrition (HLPE) only discussed forests in the context of sustainability, but did not mention their importance for people’s livelihoods or nutrition (59). Similarly, forests were only mentioned in the context of agroforestry in the HLPE report on “agroecological and other innovative approaches for sustainable agriculture and food systems” (60), which was surprising given the subject matter covered. However, there is some recognition of the benefits of forests for nutrition, such as in the HLPE report on “sustainable forestry for food security and nutrition” (61), but, on the whole, the subject is overlooked and should be better integrated into both national and international food and nutrition strategies. This might also offer potential “win–wins” in terms of meeting nutrition goals as well as conservation and environmental goals.

It is important to note that, while this study found that forest fragmentation might actually increase consumption of fruit and vegetables, our main policy recommendation is for the preservation of forests (we would not promote forest fragmentation, based on our findings). This is because the relationship between forest cover loss and diet quality in this study is causal, whereas the relationship between forest fragmentation and diet quality is not. Similarly, there is only one other study, to date, that has found an association between forest fragmentation and diets (34); thus the relationship between forest fragmentation and food consumption is still very unclear and requires further research. Ultimately, mitigating forest loss offers a clear win–win in terms of food security, biodiversity, and ecosystem services, whereas forest fragmentation does not.

Based on our findings, we identify four key directions for future research. Firstly, while our study establishes a causal relationship between deforestation and rural people’s dietary quality, more work is needed to clarify whether this relationship holds for other countries as well. Secondly, studies linking forests and diets would benefit from more detailed food consumption data that are collected at the individual level and also differentiate the source of foods, in order to more accurately capture foods that may have come from the forest. Thirdly, studies should aim to use a suite of measures to assess diet quality, including nutrient intake and adequacy levels, and quantities (grams or at least servings) of nutritionally important food groups. Lastly, future studies should consider not just forest amount but also forest configuration.

Materials and Methods

Site Selection and Household-Level Data. Tanzania was selected as an appropriate case study country to examine the impact of forest change on diets, given the high deforestation rates (62, 63), high reliance on ecosystem services (64, 65) with around 30% of the population living within 5 km of a forest (17), and high rates of hunger and malnutrition (66). According to a recent study by Doggart et al. (63), deforestation in Tanzania is largely a result of cropland expansion (i.e., the production of maize, sesame, cowpea, and sorghum), with other drivers including land clearing for livestock grazing, fuelwood collection, and charcoal production. This study used household data from a series of National Panel Surveys (NPS) from Tanzania collected as part of the LSMS. Data are freely available from the World Bank online database. We extracted data for rural households that matched across data collection waves one (2008–2009), two (2010–2011), and three (2012–2013) of the panel study. Households that remained in the same forest site across waves 1, 2, and 3 (and 4, respectively, for the parallel waves) were included in the analysis. The results were remarkably close to estimates from the original 2008–2009 data (73). The household-level data were across three waves, that is, a 5-y period in this study. While there might be a time lag of more than 5 y before the effects of deforestation on biodiversity can be fully seen, we consider a 5-y period sufficient to examine concurrent changes in forest cover and people’s diets.

Each wave of the LSMS panel data for Tanzania included a 7-d recall survey to collect data on household food consumption, where respondents reported all foods consumed by the household (and the quantities) over the 7 d preceding the interview. The survey considered 59 predefined food and drink items/categories (some foods were considered as individual items, i.e., “rice” and “eggs,” whereas others were broader categories, i.e., “peas, beans, lentils, and other pulses”). These data were used to estimate dietary diversity, consumption of nutritionally important foods; and energy, protein, and micronutrient adequacy levels. In this study, we focused on dietary diversity as an indicator of general diet quality for a household. We used the Food and Agriculture Organization (FAO) guide for measurement (70) to allocate the LSMS food items into the appropriate MDDS food groups. Households were awarded

Deforestation reduces fruit and vegetable consumption in rural Tanzania

6 of 9  Hall et al.

https://doi.org/10.1073/pnas.2112063119
one point if they had consumed at least one food from one of the MDDS food groups, and a zero if not. Thus, the score ranged from 0 to 10 and was continuous (the higher the score the better). The 10 groups considered were 1) grains, white roots and tubers, plantains; 2) pulses; 3) nuts and seeds; 4) dairy; 5) meat, poultry, and fish; 6) eggs; 7) dark green leafy vegetables; 8) other vitamin A–rich fruits and vegetables; 9) other vegetables; and 10) other fruits. A complete list of all LSMS food items included in each group is provided in Si Appendix, Table S5.

Fruit and Vegetable Consumption. We also looked specifically at fruit and vegetable consumption, given that consumption of these foods is clearly associated with positive health outcomes (73, 74), and other studies from Tanzania have shown that a key forest–diet pathway is the direct provision of wild fruits and vegetables to Tanzanians (34, 39). We calculated total fruit and vegetable consumption (grams per day) for the household as a whole, and then worked out an average consumption at the individual level using the AME approach (described in detail in the following section). We compared these individual intakes with recommended amounts of at least 400 g per person per day according to the World Health Organization (WHO) (74). We also calculated the total consumed amounts (per capita per day) of each individual fruit and vegetable category listed in the LSMS as follows: 601, “onions, tomatoes, carrots & green pepper, other vegetables”; 602, “spinach, cabbage & other green vegetables”; 603, “canned, dried and wild vegetables”; 701, “ripe bananas”, 702, “citrus fruits”, and 703, “mangoes, avocados and other fruits.”

Energy and Nutrient Adequacy. Using the 7-d recall data, we estimated total apparent intakes of dietary energy, protein, iron, zinc, and vitamin A (note that we measured vitamin A using retinol activity equivalents [RAE] in order to include vitamin A from carotenoids in plant foods) and calculated the adequacy of these intakes to meet the recommendations for each household. Food composition tables were used to estimate the energy and nutrient contents per 100 g of all LSMS food and drink items (Si Appendix, Table S4). Reported quantities of each item consumed by each household were converted from varying units (kilograms, liters, pieces, etc.) into grams using tables from Joy et al. (75). We then subtracted the inedible portions (i.e., fruit skins and peels) from the total reported weights to obtained consumed weights. Using the estimated energy and nutrient contents of each food/drink, combined with the consumed weights, we estimated total household intakes of energy and nutrients (per week and then averaged per day). To compare those household intakes of different sizes and compositions, we used the AME approach (76). The AME approach enables the estimation of inhousehold allocation of food, as it assumes that food will be distributed to each member of the household according to their individual requirements (weighted against the requirements of an 18- to 30-y-old male). Using this approach, we assigned each household member an AME value (using age and sex data obtained from the LSMS), and then summed these to get a total household AME value. Total household energy and nutrient intakes were then divided by the household AME value in order to calculate mean intakes for each household.

To estimate the adequacy of these intakes to meet requirements, we first identified the recommended intakes of energy, protein, and the three micronutrients for each individual household member, based on their age and sex and assuming standard body heights and weights. For energy, we used the AME values which weight energy requirements of different age and sex groups to that of an average 18- to 30-y-old male with moderate physical activity levels who requires 3,000 kcal d−1. Note that these values are the same as those given in the WHO report on energy and protein requirements (77). For protein, we used “safe intake” values from the WHO report (77). For iron, we used Recommended Nutrient Intakes (RNI) provided in the WHO report on human vitamin and mineral requirements (78). A 5% bioavailability factor was assumed given Tanzanian diets are low in animal sources of iron (i.e., haem iron) which has higher bioavailability than nonhaem iron found in plant foods. We assumed given Tanzanian diets are low in animal sources of iron (i.e., haem iron) which has higher bioavailability than nonhaem iron found in plant foods (78). Once recommendations were calculated for all individuals within each household, those values were summed to give a “total household recommended intake” which could then be compared with the “total household intake” of energy, protein, and each micronutrient. Methods for calculating a nutrient adequacy ratio were followed, whereby ratios were capped at one (i.e., 100%) in cases where intake exceeded the requirement (81).

Forest Cover Data. For each LSMS cluster, a georeferenced point is given, but this point has been randomly displaced by 0 km to 5 km for 99% of the clusters, for confidentiality purposes. For the remaining 1% of clusters, the random displacement of the georeferenced point is up to a maximum of 10 km. To account for this random spatial displacement as well as to capture a reasonable distance that people were likely to travel for hunting and collecting wild foods (82), we measured forest cover and configuration in a 10-km-radius circle surrounding each LSMS cluster. We used the publicly available 30-m-resolution optical tree cover dataset from 2000 to 2016 (83). Using Godot and OrthoEngine, we imported those tiles (four) covering the spatial extent of Tanzania. The imported data showed the percentage tree cover (ranging from 0 to 100) in each pixel, with trees defined as vegetation taller than 5 m. We derived tree cover in the years of the LSMS waves (2008, 2011, and 2013) by masking water, adding forest cover gain, and subtracting forest cover change from the base year 2000.

To create a forest cover map, we classified each pixel to a binary forest/no forest classification, using a “forest” threshold definition of 30%. We also tried other forest threshold definitions (10% and 60%) based on thresholds used by the FAO and United Nations Framework Convention on Climate Change (84, 85). We chose 30%, as it resulted in a forest cover map that best matched land cover maps for Tanzania. We used the landmatics R package (86) to extract percentage forest (which was accordingly transformed into hectares of forest cover) and number of forest patches (using an eight-cell rule for delineating patches), for each 10-km-radius circle in each year of the LSMS data. The correlations between forest cover and forest patches, as well as between changes in forest cover and changes in forest patches, are summarized in Si Appendix, Fig. S5.

Control Variables. A number of variables that were hypothesized to affect people’s diets and thus confound the relationship with forest change were controlled for in the analysis. We controlled for household characteristics which have been shown, in other studies, to be significant predictors of dietary quality, including household size (87, 88), the age and sex of the household head (89), and the education level of the household head (90). We calculated an asset-based wealth score to be used as a proxy for economic status. We tried other forest threshold definitions (10% and 60%) based on thresholds used by the FAO and United Nations Framework Convention on Climate Change (84, 85). We chose 30%, as it resulted in a forest cover map that best matched land cover maps for Tanzania. We used the landmatics R package (86) to extract percentage forest (which was accordingly transformed into hectares of forest cover) and number of forest patches (using an eight-cell rule for delineating patches), for each 10-km-radius circle in each year of the LSMS data. The correlations between forest cover and forest patches, as well as between changes in forest cover and changes in forest patches, are summarized in Si Appendix, Table S6. An asset-based wealth score has been shown to be less susceptible to measurement error than income data (92) and is a good proxy for economic status over time.

Given the observed linkages between farm production diversity and diet quality in smallholder farm households (44, 93), we controlled for the number of different crops cultivated by each household. Similarly, we controlled for livestock ownership (dichotomous variable), as some studies have linked livestock ownership with improved food security outcomes in rural settings (94–96). We also identified various important geographical variables shown to affect diet quality (such as road access, market access, and elevation). While these were not controlled for in the time-series models directly (as they either did not change over time or the LSMS did not track changes over time) they were included when creating the CBGPS weights (discussed in the following section). Lastly, we controlled for seasonality, given that previous studies have shown dietary differences between the rainy and dry seasons in Tanzania (in part, due to the availability of wild foods) (22, 28, 97, 98).

Statistical Analysis. We tested whether changes in forest cover (hectares) and configuration (number of forest patches) over the 5-y study period were associated with concurrent changes in dietary diversity, fruit and vegetable consumption, and energy and nutrient adequacy. Note that forest cover change was our key “treatment” variable. We also controlled for several household-level and geographical variables as stated previously. A summary of all variables included in the statistical analyses is provided in Si Appendix, Table S7. We used the following two-step modeling approach that combined two-way fixed-effects regression with CBGPS weights:

\[
Y_{it} = \alpha_i + \gamma_t + \beta X_{it} + \epsilon_{it}
\]

where \(Y_{it}\) represents the dietary quality indicators of household \(i\) in time period \(t\), and \(\alpha_i\) and \(\gamma_t\) are the unit and time fixed effects, respectively. Two-way fixed-effects models were selected, as this allowed us to control for both time and entity fixed effects; that is, this method eliminates biases from unobserved varia-

bles that change over time but are constant over entities, and controls for factors that differ across entities but are constant over time. It would not have been appropriate to use only household fixed effects, as we had only a small number of time periods (\(n = 3\) and so Nickell bias would likely have been large (99).
The two-way fixed-effects models were run first without the CBGPS weights, and then with. The weights were created and used to adjust for the nonrandom distribution (selection bias) of forest cover (and thus, forest cover change), our treatment variable of interest. The weights minimize the correlation between treatment and observable pretreatment covariates when included in regression models. Doing so reduces the dependence (endogeneity) between treatment assignment and outcome given covariates. If left untreated, it can bias the estimated effects of forest cover change on diet.

As pretreatment variables, we selected variables which likely influenced the distribution of forest cover change and diet quality. These included key biological variables: road access (distance from the household to the nearest major road), elevation, slope, mean annual precipitation, and number of forest patches within 10-km-radii circles, and the same socioeconomic variables used in the fixed-effects models: household size, age and sex of the household head, wealth score, highest education level of the household head, crop, and livestock ownership. Cropland expansion, a key driver of deforestation in Tanzania (63), was not included as a covariate, due to its lack of association with forest cover change (SI Appendix, section C). Possible important unobserved confounders are fuelwood consumption and charcoal production (63); however, a semiformal test of endogeneity found no evidence of the presence of unobserved confounders (see SI Appendix, section D for more information).

As the names suggest, “pretreatment” variables should come before the treatment variable (forest cover), to be sure that they themselves have not been influenced by the treatment. For instance, wealth and crop and livestock counts could have been influenced by forest cover. To reduce the likelihood of this, we used the earliest data available for all socioeconomic variables (i.e., wave one, 2008–2009) and found little evidence to suggest that they have been influenced by forest cover (both their correlation with forest cover and their contribution in explaining forest cover [partial r^2] was very low).

The CBGPS method builds on popular propensity score methods applicable only for binary treatments (100, 101). In addition to being applicable to a continuous treatment variable such as forest cover change, the CBGPS method is also found to be more robust to model misspecifications (102). In effect, the method mimics the experimental condition of randomness which allows for causal inferences to be made. We used the CBGPS method’s parametric approach from the CBGPS package in R. This generated weights with acceptable low correlation levels (all weighted correlations to forest cover were below 0.2, compared to an original maximum correlation of 0.42), despite being far less computationally intensive than the available nonparametric approach. These weights were then included in two-way fixed-effects linear regression models which were carried out using the plm package in R.

Data Availability. Food consumption and all other socioeconomic data are publicly available from the World Bank’s microdata library (https://microwdata.worldbank.org/index.php/catalog). Tree cover data are publicly available on Global Forest Watch’s Open Data Portal (http://data.globalforestwatch.org/). All other relevant data are available in the article and/or SI Appendix.

ACKNOWLEDGMENTS. C.M.H. and L.V.R. were funded by the European Research Council under the European Union’s Horizon 2020 Research and Innovation Programme (Grant Agreement 853222 FORESTDIET). We thank the editor and anonymous reviewers for very helpful comments.
