EthioChicken Impact Evaluation:

Understanding income and nutritional effects of expanding access to high productivity chickens for Ethiopian smallholder farmers in the early days of ownership

Endline Report; June 29, 2018

Prepared by IDinsight

This report was prepared by IDinsight in partnership with EthioChicken and Acumen, with funding from the Bill & Melinda Gates Foundation.

Executive Summary

EthioChicken is a private company that breeds high-productivity chickens and sells them to smallholder farmers in Ethiopia. EthioChicken distributes chickens through its network of agents (who in Oromia grow chickens from one day old to 56 days before distributing) and government agricultural extension agents who take orders and deliver chickens to smallholder farmers.

EthioChicken makes two breeds available: Bovans layers and Sasso dual-purpose birds. Both breeds are expected to grow four times faster and produce four times as many eggs as traditional Ethiopian backyard birds,[1] while also being able to grow and develop in the local environment. These higher-productivity birds have the potential to increase household income and provide sources of animal protein for consumption. As women are primarily responsible for raising chickens in Ethiopia, higher productivity chickens could also lead to increased decision-making power for women.

This report presents the results of an impact evaluation of providing access to improved chickens, leveraging EthioChicken's expansion into the state of Oromia in Ethiopia. It studies the short-term impact of owning improved chickens on smallholders' income, nutrition, and female decision-making roughly 6-14 months after purchasing improved chickens. To identify the causal effect of owning improved chickens, we match purchasers of improved chickens with households who did not have access to these chickens. The matching algorithm ensures comparability on a number of key baseline characteristics as well as a predicted probability of purchasing chickens (generated using a machine learning model). Our endline data, collected from October 2017-February 2018, was collected 6-14 months after chickens became available in our treatment areas. Since EthioChicken's chicks are sold at 56 days old in Oromia and ferenj chickens take around 5-6 months to begin laying eggs, the timing of the endline assures that the chickens have had the opportunity to mature and start producing eggs or be sold for meat.

We find that households who purchased improved chickens:[2]

- Are likely to keep rearing ferenj chickens, as around 65% have improved chickens at endline. Purchasers seem to rear ferenj chickens *in addition to* rather than *in place of* habesha chickens. Both the treatment and control groups of this study had an average of 2.5 habesha chickens, while the treatment households also had an average of 2.2 ferenj chickens.
- **Produce around 6.6 extra eggs per week** (compared to 4.5 eggs/week in the comparison group) and **consume an additional 3.0 eggs per week** (compared to 3.6 eggs/week in the comparison group). Purchasers of improved chickens consumed eggs on an additional 0.7 days per week (compared to 0.80 days per week in the comparison group).
- Increase sales of eggs, resulting in an **increase in average income from egg sales in the prior month of around 14 ETB (.51 USD)**. This is compared to average income from eggs of around 7 ETB (.26 USD) in the comparison group.
- Increase their income from chicken sales. Revenue **from chicken sales in the last 6 months increased by around 88 ETB (3.21 USD)** over an average of around 103.5 ETB (3.76 USD) in the comparison group.

---

[1] Aglionby, John. "EthioChicken: Ethiopia's well-hatched idea." *Financial Times*. Nikkei Inc., 16 March 2018.

[2] All results stated in the executive summary are statistically significant at the 5% level except for the result on women's egg consumption, which is significant at the 10% level.

- **Increase monthly expenditure on chicken rearing by around 28 ETB (1.02 USD),** compared to average expenditures of 7 ETB (.29 USD) in the control group. The main expenditure category for owners of improved chickens was feed. The increase in average expenditures seen in treatment households is a consequence of a sizeable minority of households taking on greater expenses; the median treatment household saw a marginal increase in expenses compared to the control group.
- Increase egg consumption. Households have an **increased Food Consumption Score** driven by increases in egg consumption. Based on 24-hr dietary recall, we find that **women are 4 percentage points more likely to consume eggs** (compared to an average of 4.9% in the control group), but not more likely to consume meat. We find that **children under 5 are not significantly more likely to consume eggs** or animal protein.
- Do not see an improvement in women's decision-making power over how income from chickens is used, but report modest gains on a constructed index of women's decision-making that also takes into account participation in and input on certain chicken-related activities.

Additionally, **we do not find any evidence of significant spillovers** onto households who did not purchase chickens but lived near other households who did purchase chickens. This is not surprising given the modest magnitude of direct effects that were hypothesized to have driven spillovers, specifically egg production.

Overall, we find that purchasing improved chickens results in a statistically significant increase in income from egg and chicken sales compared to households that own only local breeds of chicken. However, the impact on households' well-being is likely to be limited given the small magnitude of the effects found; more transformative impact might require owning larger flocks of improved chickens. Furthermore, households experience higher income related to poultry, but also higher poultry-related expenses. Although we do not see increased average net poultry income in our sample, there is significant heterogeneity in the results and the impact of improved chicken ownership on long-term profit remains unclear.

Lastly, we find rather modest results on nutrition and no improvements on children's consumption of animal protein. This may be partially explained by issues of power: since the overall increase in egg consumption by families is quite modest, it is hard to tease out increases for individual family members. However, it may also point to the need for complimentary interventions if chicken rearing is going to reach its potential as a nutrition booster for children.

Recommendations for future research:

- Assessing the impact of high productivity chickens on medium to long run outcomes to facilitate a better understanding of nutrition and income effects over the lifecycle of chicken ownership
- Exploring why chicken expenses were higher than expected, and whether productivity gains can be achieved without these extra expenses
- Understanding the evolution of chicken expenses, income, and profits as flock size increases
- Addressing the overall income effects of chicken rearing and whether investing in chickens is in addition to or in substitution of other potential investments
- Determining whether complementary interventions aimed at women's nutrition, women's empowerment, and children's nutrition can boost the effect of high productivity chickens on those outcomes

Table of Contents

List of Figures

List of Tables

Introduction

Motivation

Many in the development space see village poultry as a potential tool for smallholder income generation and improved animal protein consumption.[3] In rural Ethiopia, chickens have the potential to be a tool for savings as well as an income-generating instrument for smallholder farmers. A study conducted in Southern Nations, Nationalities and People's Regions (SNNPR) in Ethiopia estimates that poultry contributes between 2-10% of a household's income.[4] In addition to its potential to contribute towards a household's income, poultry rearing can increase a household's access to and consumption of animal-sourced foods, which are a source of protein and micronutrients such as iron, zinc and vitamin A[5]. The potential benefit to children is critical in Ethiopia where 38.4% of children under the age of two are stunted (low height-for-age), an indicator that reflects the cumulative effect of chronic malnutrition.[6] Furthermore, home poultry rearing has also been seen as a potential tool to empower women, though it is still unclear if growing smallholder poultry business will actually be a lever for empowerment.[7] Donors and researchers have shown interest in understanding the impacts of poultry rearing, particularly disaggregated by gender.[8]

Additionally, foreign chicken breeds are increasing in popularity in Ethiopia and throughout sub-Saharan Africa with potential to provide greater benefit than indigenous alternatives. Studies have shown that some foreign and crossbred chickens have higher body weight and produce more eggs.[9] This could lead to increased consumption of animal protein and additional income for households. However, little empirical research has been done to quantify the ultimate impacts of households rearing foreign chickens on their nutrition and income. While some studies have estimated the benefits of foreign chickens in terms of egg production, egg quality, and chicken size in Ethiopia,[10] there is limited research into the impact of rearing foreign chickens on ultimate outcomes like nutrition and income.

As governments, companies, and donors consider additional investment in providing access to improved chickens in Ethiopia, it is necessary for stakeholders to understand how poultry rearing impacts

---

[3] For instance, this blog post by Bill Gates: Gates, Bill. "Why I Would Raise Chickens." *gatesnotes*. The Gates Notes LLC, 7 June 2016.

[4] Bush, J. *The threat of avian flu predicted impacts on rural livelihoods in Southern Nation, Nationalities and Peoples Region (SNNPR), Ethiopia.* Wolfville: Food Economy Group, 2006.

[5] Kryger, K.N. et al. *Smallholder Poultry Production – Livelihoods, Food Security, and Sociocultural Significance.* Rome: Food and Agriculture Organization of the United Nations, 2010. Print. FAO Smallholder Poultry Production No. 4.

[6] Central Statistical Agency (CSA) [Ethiopia] and ICF. *Ethiopia: Demographic and Health Survey.* Addis Ababa, Ethiopia and Rockville, Maryland: CSA and ICF, 2016.

[7] Guèye, E.F. "The Role of Family Poultry in Poverty Alleviation, Food Security and the Promotion of Gender Equality in Rural Africa." *Outlook on Agriculture.* 29.2 (2010): 129-136. Web. 22 March 2018.

[8] Guèye, E.F. "Gender issues in family poultry production systems in low-income food-deficit countries." *American Journal of Alternative Agriculture.* 18.4 (2003) 185-195.

[9] Tadesse, D. et al. "Study on productive performances and egg quality traits of exotic chickens under village production systems in East Shewa, Ethiopia." *African Journal of Agricultural Research.* 8.13 (2013) 1123-1128.

[10] Tadesse, D. et al. "Comparative study on some egg quality traits of exotic chickens in different production systems in East Shewa, Ethiopia." *African Journal of Agricultural Research.* 10.9 (2015) 1016-1021.

households and individuals within the household. To shed light on this question, IDinsight conducted an impact evaluation on the breed and distribution model propagated by EthioChicken, a private, commercial-scale hatchery in Ethiopia.

Background on EthioChicken

EthioChicken is a poultry company based in Addis Ababa, Ethiopia specializing in the hatching and sale of dual purpose and layer chickens to smallholder farmers in Ethiopia. Compared to indigenous breeds (known as "habesha" chickens), which are estimated to produce 40-60 eggs per year,[11] specialized dual purpose and layer chickens (referred henceforth to as "ferenj", or exotic breeds of chickens) are specially bred to grow larger and lay more eggs for a longer period of time. The ferenj chickens produced by EthioChicken include male and female dual purpose (Sasso breed) and female layers (Bovans breed). The female ferenj chickens can produce eggs earlier in life and in greater number than habesha chickens. According to a survey of poultry raisers in the SNPPR regions of Ethiopia, Sassos begin laying after an average of 5.9 months and Bovans begin laying after an average of 5.7 months, while habesha chickens do not begin laying until after an average of 7.1 months.[12] Ferenj chickens tend to lay more consistently: Sassos reportedly produce an average of 16.2 eggs per month and Bovans produce an average of 22.2 eggs per month, while habesha chickens produce an average of 12.6 eggs per month.[13] These levels of productivity are close to those seen in the evaluation data—among households in the final sample, ferenj chickens laid an average of 18.5 eggs per month while habesha chickens laid an average of 10.8 eggs per month. Rearing ferenj chickens may thus give farmers access to a more consistent stream of income from eggs and increased income from the sale of chickens, while increasing access to eggs for consumption.

EthioChicken works through regional governments to sell and deliver day-old-chicks (DOCs) to independent entrepreneurs, known as agents, based in the woreda.[14] Agents are trained in chicken nutrition and health, and raise the chickens to 45 days old in Amhara, SNNPR, and Tigray and to 56 days old in Oromia before selling them on to farmers.[15] Developmental Agents (DAs), who are agricultural extension officers within kebeles, collect chicken purchase orders from farmers and then deliver old chickens from the agents to farmers.

EthioChicken began selling chickens in the Tigray region in 2010 and has since grown to the Amhara, Oromia, and Southern Nations, Nationalities, and Peoples' (SNNPR) regions. Acumen's Lean Data surveys have found that chicken-owning households in SNNPR own an average of 6 ferenj chickens and 3 habesha chickens, with a median length of ferenj ownership of 2 years. In Tigray, chicken-owning households own an average of 10 ferenj chickens and 3 habesha chickens, with a median ownership length of 3.4 years.

---

[11] Alemu, D. et al. *Overview and background paper on Ethiopia's poultry sector: Relevance for HPAI research in Ethiopia.* Washington, D.C.: IFPRI, 2010.
[12] Getiso, A. et al. "Production performance of Sasso (distributed by EthioChicken private poultry farms) and Bovans brown chickens breed under village production system in three agro-ecologies of Southern Nations, Nationalities and Peoples' Regional State (SNNPR), Ethiopia." *International Journal of Livestock Production.* 8.9 (2017) 145-157.
[13] *Ibid.*
[14] Administrative zones in Ethiopia are grouped into regions, zones, woredas and kebeles; kebeles are the lowest formal administrative unit.
[15] Chickens raised for distribution in Oromia are kept until 56 days old because the disease environment in Oromia requires them to receive an additional vaccine.

With support from regional and federal governments, EthioChicken has been able to expand operations in these regions, with production reaching 10 million chicks in 2017 from five hatcheries across the country.[16]

At the inception of our study, EthioChicken was planning an aggressive expansion into Oromia, Ethiopia's largest region by land and population size. Due to EthioChicken's lower penetration rates in this region, Oromia offered the opportunity to leverage EthioChicken's natural expansion plan for an impact evaluation to measure the impact of ferenj chicken uptake on smallholder farmers. EthioChicken planned to expand aggressively into Oromia in late 2016 and through 2017. As a result of the endline data collection taking place not long after EthioChicken's expansion—allowing for less ownership time than the median length in SNNPR or Tigray—the results from this evaluation should be interpreted as the short to medium term effects of purchasing ferenj chickens.

For EthioChicken, the key first step to expanding into a new territory is recruiting agents. EthioChicken made an arrangement with the ministry of youth employment in Oromia to identify potential agents and provide seed funding for them. Unfortunately, this process moved more slowly than anticipated and therefore EthioChicken was not able to recruit agents as quickly as originally hoped. However, despite the lack of local agents in many regions, local DAs were creative in sourcing chickens by purchasing them from other woredas or even other states to distribute in their kebeles. Therefore many farmers in our baseline sample had access to ferenj chickens, though the total amount of purchasers was lower than expected. It is possible that some farmers in our sample were unable to purchase as many chickens as they would have liked due to supply restrictions.

EthioChicken Theory-of-Change

We examine the process by which households are impacted by their purchase of ferenj chickens from EthioChicken using a stylized theory-of-change, shown in Figure 1. Each node connects to the following node as long as a set of assumptions holds, forming impact channels that have potential to achieve the ultimate outcomes of improved nutrition and income.

[16] Aglionby, John. "EthioChicken: Ethiopia's well-hatched idea." *Financial Times*. Nikkei Inc., 16 March 2018.

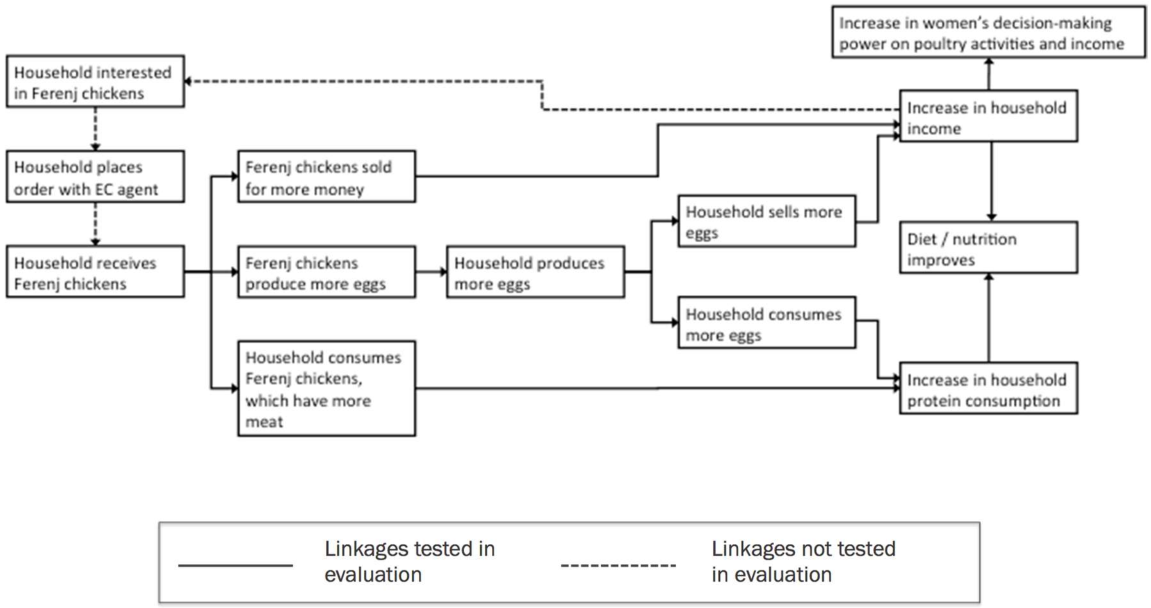

Figure 1: EthioChicken intervention theory of change

In short, households with improved chickens can expect these chickens to grow larger and produce more eggs. This may result in households eating more chicken meat and eggs, which may improve the nutrition of household members. Additionally, households can gain increased income from selling eggs and chickens. This may also improve nutrition if households use this extra income to purchase nutritious foods.

Additionally, we are also interested in the impact of chicken rearing on women's empowerment. Women are traditionally seen as the caretakers of chickens in Ethiopia. If chickens become a growing enterprise and a larger share of the household's income, women's decision-making power over poultry activities and income may positively affect their status and bargaining power in the household.

Evaluation Methodology

Study overview

This impact evaluation creates a counterfactual by matching purchasers of improved chickens to potential customers who did not have access.[17] This approach leverages EthioChicken's natural expansion plans, selecting ferenj purchasers in kebeles where farmers have been given access to ferenj chickens and comparing them to kebeles in which access to improved chickens was limited. Generally speaking, 'limited' access was defined as fewer than 2 households per kebele with ferenj chickens. We use a matching algorithm to choose comparison households that are similar on key baseline characteristics. We then construct our impact estimates by comparing key outcome indicators at endline between ferenj purchasers ("treatment") and comparison households. We also measure spillover effects by comparing outcomes between people who did not purchase ferenj chickens but live near people who purchased and matched comparisons of non-purchasers who did not live near purchasing households.

The study's baseline survey was conducted beginning in October 2016 and the endline was completed in February 2018. Major steps of the research process are outlined in Figure 2.

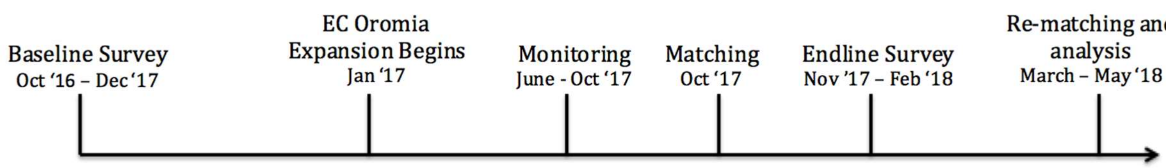

Figure 2: Timeline of IDinsight's research activities

Study Area and Baseline Sample

This study was conducted in Oromia, a state in the south-west of Ethiopia. Households in our combined direct effect and spillover samples in Oromia own an average of 3.2 chickens compared to a countrywide average of 1.94 chickens.[18] The average score on the wealth index among households in Oromia is 3.08 while the average score on the wealth index in Ethiopia is 3.09.[19] Oromia was chosen because EthioChicken had limited presence in the state at the start of the study but had rapid expansion plans for 2017. This made it feasible to leverage EthioChicken's natural expansion to create a counterfactual. In order to further narrow down the study area, IDinsight engaged with EthioChicken personnel and local government officials to understand EthioChicken's current coverage throughout the state and identify regions that were likely to gain access to chickens in 2017. In the end, IDinsight selected 153 kebeles in 7 woredas (Digeluna Tijo, Tenna, Meta, Goro Gutu, Doba, Gumay, and Kersa) to be part of the study.[20]

---

[17] A counterfactual is a measurement of what would have happened had the households not purchased chickens. Initial considerations to conduct a randomized control trial were ruled due to the fact that EthioChicken did not want to be constrained in their options for expansion.

[18] Central Statistical Agency (CSA) [Ethiopia] and ICF. *Ethiopia: Demographic and Health Survey.* Addis Ababa, Ethiopia and Rockville, Maryland: CSA and ICF, 2016. Web. Accessed: 8 May 2018.

[19] *Ibid.* The wealth index is a measure of a household's standard of living, "calculated using…data on a household's ownership of selected assets, …materials used for housing construction; and types of water access and sanitation facilities."

[20] Since the beginning of this study, Goro Gutu has been split into two woredas: Goro Gutu and Goro Muti. This study included households from both woredas.

We conducted a baseline survey to get a better understanding of our target population as well as to gather the data we would use for matching. First, we had to determine the household-level sampling frame for data collection. As the final analysis was to be conducted based on purchasers of chickens, we sought to obtain a sampling frame of people who were likely to purchase chickens. To accomplish this, we leveraged the knowledge of local DAs, who had experience working with households in their kebele. Ahead of data collection, IDinsight collaborated with DAs to collect a list from each kebele of up to 40 households who would be interested and able to purchase EthioChicken chickens once EthioChicken became available at their woreda ("would-be-buyers"). Information on types of chickens currently owned by farmers, if any, was also collected during this listing process.

The lists from the DAs contained mostly households who already owned habesha chickens and some households who were already rearing ferenj chickens (likely from sources other than EthioChicken). In order to get a mix of respondents, we stratified based on current chicken ownership, and randomly chose 20 households from each kebele. Households who already owned ferenj chickens were given lowest priority for sampling, and therefore were only sampled if there were not enough households who did not own ferenj chickens on the list. Overall, our baseline sample consisted of 25% households with no chickens, 64% households with just habesha chickens, and 11% households with ferenj chickens.

The baseline was conducted by IDinsight between September 2016 and December 2016. Enumerators explained the scope of the evaluation to respondents, who provided written consent to participate in the study. Enumerators were hired by IDinsight and had no information about EthioChicken or its expansion plans. Whenever possible, the member of the household who was primarily responsible for the daily care of chickens was interviewed for the survey. When that person was not available to answer, another attempt was made at a later time in the day. If the person was still not available on the second attempt, the head of the household was selected to be interviewed. We were able to survey the person responsible for chickens 98.5% of the time. Of these respondents, 92.9% were women. If the respondent was a female, we asked her the women's nutrition questions. If not, we asked a female household member to respond. For children's nutrition questions, we asked about the eldest child under the age of 2 (if available) or otherwise the youngest child under 5.[21] These questions were also asked to a woman in the household who had knowledge of the selected child's feeding practices. In total, 3,153 households were surveyed at baseline.

Data was collected on the electronic data platform SurveyCTO and backed up daily onto IDinsight computers and an encrypted cloud server. Data quality was monitored in real-time, with additional thorough weekly data checks. Enumerators who had unusually long or short survey times were given additional training and were monitored in the field. Back-checks[22] were also conducted to ensure that data was collected accurately. Managers coordinated logistics of surveying and relationships with local government officials. Data Quality Supervisors conducted back-checks to randomly confirm data from households and coach enumerators on data collection techniques.

---

[21] We prioritized children under two since this is the age in which nutritional changes are expected to have the most long-term impact. However, very young children were not prioritized since they are unlikely to eat whole foods (like eggs or meat) and are likely to obtain their protein requirements elsewhere (breast milk).

[22] A subset of households was randomly sampled and a dedicated enumerator called to confirm a small subset of survey questions.

The IDinsight team obtained permission from the Federal Ministry of Livestock and Fisheries, the Oromia Livestock Bureau, and local government offices to conduct a household survey. Additional written permission was obtained from each woreda and kebele included in the study. Team supervisors were tasked with data quality, including random sit-ins during enumerator interviews.

Monitoring

To determine who from our sampling frame purchased ferenj chickens, we conducted a monitoring exercise 6-10 months after baseline. This began with contacting all DAs to try to get an idea of whether any chickens had been distributed in their kebele. We followed this up with phone calls to all of our respondents who had provided phone numbers, inquiring if they knew about ferenj chickens being available and if they had purchased ferenj chickens since baseline. We were able to speak to around 50% of our sample over the phone. Then, in areas where we determined that some chicken sales had likely taken place, we visited the rest of our sample in-person to verify chicken purchases. In places the DA claimed that no chickens were distributed and no contacted households reported purchasing chickens, we simply assumed that household we didn't contact also hadn't purchased chickens.[23] Of the 3,153 households surveyed at baseline, we were able to get in contact with 2,669 of them during monitoring, covering all of the woredas sampled at baseline.[24] Of these households, 746 reported purchasing ferenj chickens since baseline.

One important caveat is that although EthioChicken was the largest supplier of ferenj chickens in Oromia during our study time, due to previously-discussed expansion constraints sometimes DAs procured ferenj chickens from different sources. Many of the chickens from other sources are the same or similar breeds to those promoted by EthioChicken and farmers do not know the difference. Therefore, in this study we consider all famers who purchased any type of ferenj chickens as potential to be "treatment", even though some chickens likely came from sources other than EthioChicken.

Treatment and Control Assignment

After determining individual ownership status, we then separated kebeles into treatment and control. Ideally, a treatment kebele is one in which ferenj chickens were offered for purchase, while a control kebele is one in which households did not have access. Kebele level treatment assignment was determined by considering two sources of information from the monitoring data: the number of households in the kebele that reported purchasing EthioChicken chickens as well as those who reported having access to ferenj chickens. An ideal control kebele would have no chicken purchasers as well as every household reporting that they didn't have access to chickens. An ideal treatment kebele would be indicated by all households reporting that they were offered ferenj chickens, with many choosing to purchase.

In reality the data was much messier, with many villages having just 1 or 2 ferenj chicken purchasers, and reported access to chickens within a kebele exhibited a lot of variation. Our understanding was that in many kebeles, the DAs delivered chickens to their friends without taking orders more widely, which could explain the instances of limited but not non-zero access. Also, our enumerators explained that it was very

---

[23] We verified chicken ownership at endline. If a household that we had determined was a potential control turned out to have purchased ferenj chickens, they were removed from the sample prior to rematching.

[24] If we were unable to determine whether a household had purchased chickens, they were dropped from the potential matching sample.

difficult for households to remember specifically if DAs had informed them about chickens being available. Therefore, we allowed some kebeles to be classified as control if they had a small amount of purchasers. Kebeles were classified as control if they had fewer than two purchasers (out of up to 20 households contacted) and the majority of non-purchasers claimed they did not have access to ferenj chickens.[25] All other kebeles were classified as treatment. This resulted in 82 treatment kebeles, and 70 control kebeles.

Household treatment assignment was determined using both kebele-level treatment status as well as household-level chicken purchasing. Households were considered potential treatments if they lived in treatment kebeles and purchased ferenj chickens since baseline. Households were considered potential control if they lived in control kebeles and had not purchased ferenj chickens since baseline. This resulted in 497 potential treatment households, and 877 potential controls.

We also selected some households in the analysis to study spillovers. To do this, we selected households as potential spillover "treatments" if they did not purchase chickens but lived in kebeles where more than seven households purchased chickens (which corresponds to around 1/3 of the sample). All control households were potential matches for the spillover comparison. This resulted in 207 potential spillover treatment households and 877 potential controls. Although we used this simple heuristic to do sampling for the spillover sample, as described later we implement a more sophisticated approach during the analysis phase.

Matching

We then matched potential treatment households to potential controls to obtain a study sample that was balanced on baseline values of key variables. The variables we use are (a) a prediction of purchasing chicken from a machine learning model and (b) our primary outcome variables.

The first step in the matching was to calibrate a model that would predict which households would purchase chickens. To do this we turn to the machine learning literature and use our treatment villages to fit a predictive model of chicken purchasing.[26] We use a common technique called a LASSO regression, which is popular in the predictive literature for being simple to use and performing well out of sample. We train and cross-validate this model using our treatment kebeles and the entire set of variables gathered at baseline, and then use the model to generate probabilities of purchasing for our entire sample. Each household in our sample is assigned a probability of purchasing from the model, and this probability of purchasing becomes variable upon which we conduct the matching. Although the final predictive model uses 40 variables from baseline, one key predictor of purchasing ferenj chickens was current ownership of ferenj chickens.[27]

---

[25] There were two exceptions to this rule, due to minor classification errors. Shoa Totobi Kebele in Kersa Woreda was classified as control although it had two purchasers and 60% of other households claimed to have access to chickens. Lole Habosare Kebele in Digulena Tijo woreda was classified as treatment although it has two purchasers and 38% of other households reported having access.

[26] Lara Ibarra, Gabriel; Mckenzie, David; and Ruiz Ortega, Claudia. *Learning the impact of financial education when take-up is low.* Working Paper 8238, Impact Evaluation Series. Washington, D.C.: World Bank Group.

[27] Some previous runs of the predictive model showed ferenj ownership as the *only* important variable, so we decided to do exact matching on this variable.

In addition to the probability of purchasing, we also wanted to achieve balance on baseline values for our key outcome variables and a few other variables that we thought might affect outcomes.[28] There variables were: a dummy for owning ferenj chickens, a dummy for having owned ferenj chickens for more than one year, the total number of ferenj chickens, the total number of chickens, income reported from selling chickens, income from eggs, the number of eggs eaten by the household in the last 7 days, and a dummy for whether the woman consumed protein in the last 24 hours.

The next step is to actually conduct the matching. We used a technique knows as "genetic matching,"[29] which is an iterative approach that attempts to get as good of a match as possible simultaneously on all variables. The matching gave us balance on key variables and generated an initial sample of 370 treatment and 370 control households for the measurement of the direct effect, and 186 treatment (households in treatment kebeles that did not purchase) and 186 control households (households in control kebeles that did not purchase) to estimate potential spillover effects.[30]

Additional details on the machine learning and matching algorithms are given in the Technical Appendix.

Endline Sample

At endline we attempted to revisit the treatment and control households designated by our initial match (1011 households) and we also included a set of alternative household matches in anticipation of attrition and disagreements in chicken ownership status between monitoring and endline. To make sure we had a large enough sample in spite of potential unforeseen problems, we interviewed 293 alternative households that were close but not final matches, conducted in cases where our team was already in a kebele and had excess manpower to interview additional households in that area. In the end, we surveyed 1,209 households at follow-up, of which 937 were from the initial sample and 272 were potential replacements. All permissions and data quality checks performed at baseline were also repeated during the endline survey.

Due to civil unrest in the Oromia region during 2017 and 2018, surveying was frequently halted to protect the security of our data collection teams. An extreme case occurred during endline data collection in December 2017 when protest in East and West Hararghe halted data collection for over a month. The timing of the postponement of the endline survey happened during the key Christian holidays on the Ethiopian calendar, separating data into pre and post-Christmas. In our analysis, we control for these timing issues in expectation that there may be some differences in consumption patterns between these time periods.

---

[28] We added these other variables because just matching on the predicted probability did not lead to good baseline balance on our primary outcome indicators. Most additional matching variables are primary outcome variables. The one exception (a dummy for having owned ferenj chickens for more than one year) had proven important in certain variations of the prediction model, so we felt that it was especially important to have balance on it as well.

[29] Diamond, Alexis and Sekhon, Jasjeet. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *The Review of Economics and Statistics*. 95.3 (2013) 932-945.

[30] Some control matches for the spillover were the same as the control matches for the direct effect.

Re-matching

Due to attrition and disagreements between monitoring and endline, we conducted a re-matching of our endline sample. Once endline data collection was complete, the same matching methodology (using baseline data) was used to rematch households within the pool of treatment and controls we were able to reach at endline.[31] Figure 3 gives a picture of the evolution of sample size in this evaluation. Over the course of data collection, we moved from a baseline of 3,153 households to a final combined direct effect and spillover sample of 849. Table 1 shows the breakdown of this final sample into treatment, control, spillover treatment, and spillover control categories. Table 1 also shows the sample of households used for the infant and young child feeding outcomes, which is a subset of the main sample.

Figure 3: Diagram of household sample evolution

	Direct Effect Sample	Spillover Sample	Child-level Outcomes Match
Treatment Households	307	175 (33 of which overlap with direct effect control)	134
Control Households	307	175 (82 of which overlap with direct effect control)	134
Total Households	614	350	268
Comments	Final Main Sample = 849 (Accounting for overlap between direct effect and spillover)		This sample is a subset of the main sample

Table 1: Direct Effect, Spillover, and Child-level Outcome Samples

Each treatment household purchased a ferenj chicken between baseline and endline, but not all treatment households still owned their ferenj chickens at the time of endline data collection. Table 2 breaks down the ownership status of treatment households at endline.

Ownership Category	Number of Treatment Households
Owned Ferenj Chickens at Endline	199
Owned Ferenj Chickens within 6 Months of Endline	20
Owned Ferenj Chickens Longer Ago than 6 Months, but Purchased Ferenj after Baseline	88
Total	307

Table 2: Ferenj Ownership Status of Treatment Households

[31] We also dropped 6 households from the endline sample prior to re-matching, due to poor quality data. These households reported egg production far outside biological possibility.

For child-level outcomes, we used a slightly more complicated approach. Our survey asked children's nutrition questions for households that had children aged 6 months to five years (of the 1,209 households we surveyed at endline, 655 households had children in this age bracket). We wanted to compare treatment and comparison households who had children of similar ages at endline and were also balanced on children's baseline characteristics. However, not all households who reported children's nutrition at endline did so at baseline. For all households with children at endline we matched on all the standard baseline matching variables plus the child's age at endline. For households with children at endline who also reported child nutrition at baseline, we also matched on child's egg and protein consumption at baseline.

For spillovers, we implemented a more sophisticated approach during the re-matching phase. In order to more accurately determine which households might have been exposed to potential spillover effects, we create an exposure metric that takes into account the number of households with ferenj chickens that live nearby. For each non-purchaser of ferenj chickens, we create an exposure metric that is a weighted sum of the number of purchaser households within 1km, 5km, and 10km. Households within 1km are given a weight of 1, those within 5km are given a weight of 1/5, and those within 10km are given a weight of 1/10. The proximity rings are based on the idea that closer households should have a stronger spillover effect.

Potential spillover treatment households are those with an exposure metric in the top 25 of the distribution of exposure, while potential spillover control households are those in the bottom 25% of the distribution. Using the same algorithm as was used for the direct effect matching, we then match the potential spillover households so they are balanced on key baseline characteristics.

After re-matching our endline sample (which contained 420 possible treatment and 454 possible control households), we end up with a final sample of 307 treatment and 307 control households for measuring direct effects on household-level indicators. We have 134 treatment and control households for measuring child-level outcomes and 175 treatment and controls households for spillover effect estimation. The smaller sample sizes for child-level outcomes and spillovers means that estimates are made with reduced precision.

Figure 4 shows pre and post-match balance between treatment and control households along the key matching variables. The "distance" metric is an overall measure of closeness, akin to a propensity score, generated by the genetic matching algorithm. This figure shows the mean difference between treatment and control households—the sample is well matched on a variable if the post-match difference between treatment and control (indicated by the blue diamonds) is close to zero. All variables except binary indicators have been standardized to maintain the same range across all variables.[32] The pre-match differences are calculated from the 1376 households that were initially considered possible candidates for the direct effect comparison. The post-match differences are calculated from the 307 treatment and 307 control households identified by the final matching algorithm. On all ten inputs to the matching algorithm (including the distance metric), there is better post-match balance than there was prior to matching (pre-match differences denoted by gray squares). There is approximately a difference of zero between treatment and control along eight of the input variables; "Distance" and "Predicted Value" are

[32] The binary indicator variables are "Current FJ Owner" (whether the household owns a ferenj chicken), "Owned FJ One year" (whether the household has owned a ferenj chicken for at least one year), and "WDD Protein" (whether a woman in the household has eaten protein within the last 24 hours). All other variables are continuous and standardized in the match balance graphic.

slightly worse not because they don't have good balance on the raw data, but because their standard deviations are very narrow and standardization inflates the actual difference. To the extent that the inputs to the matching process were chosen correctly (which we believe is the case), the balance obtained by the matching algorithm provides strong support for the control households to be valid counterfactuals for the treatment households.

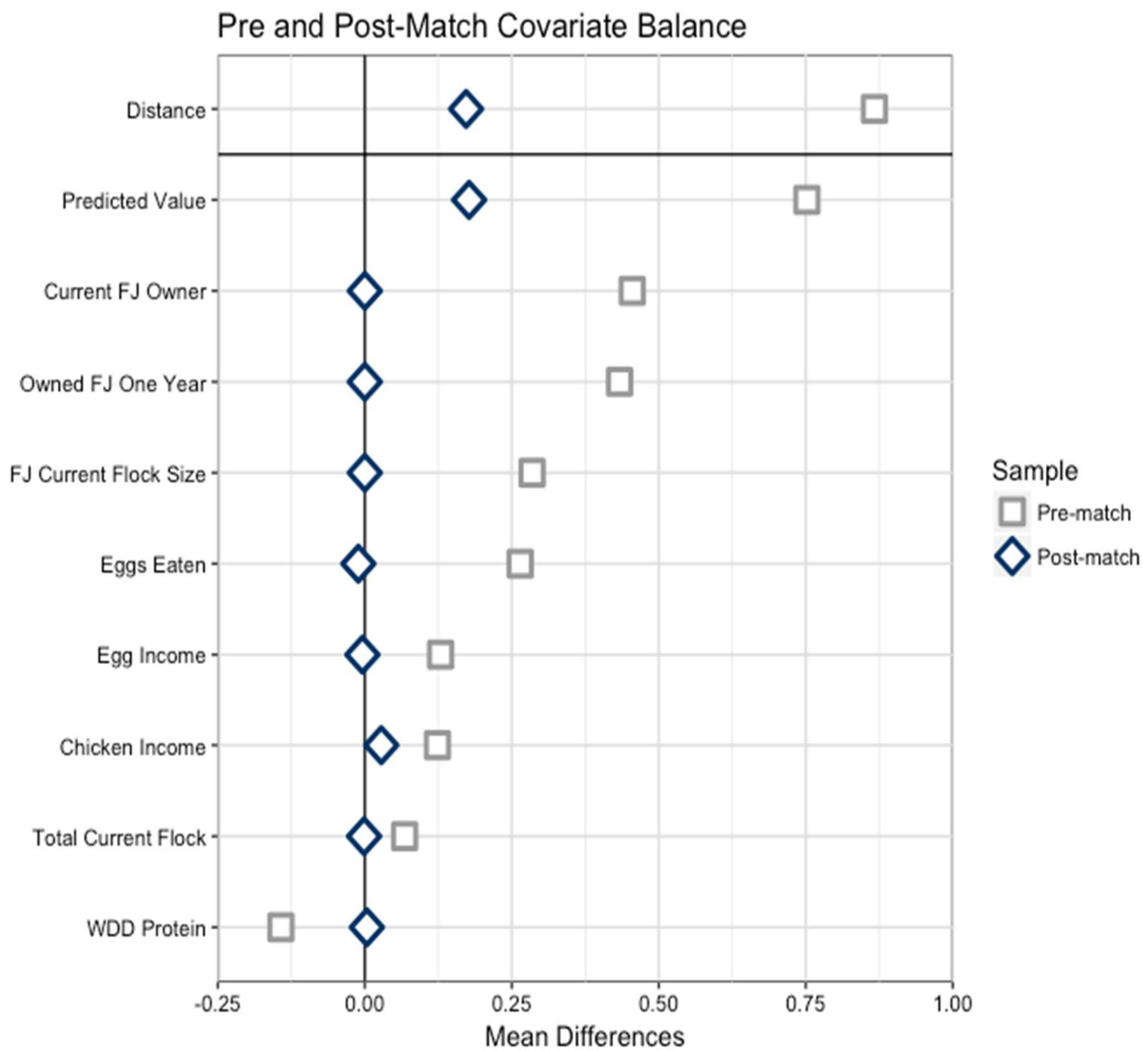

Figure 4: Pre and Post Matching Balance

We provide descriptive statistics (calculated at baseline) of the households in our sample in Table 3. We find decent balance between treatment and control even among baseline variables not included in the matching algorithm, though there is are some exceptions. Heads of household in the treatment group are more likely to have completed primary school, and control households have more goats.

Table 4 provides the average household size and the average chicken flock size of the treatment and control sample for each woreda.

	All Endline Households	Treatment Group	Control Group	P-value of Difference[33]
Household Size	6.232 (2.140)	6.225 (2.176)	6.244 (2.070)	0.924
Proportion Who Completed Primary School	0.503 (0.500)	0.544 (0.499)	0.450 (0.498)	0.027
Proportion Who Completed Secondary School	0.065 (0.247)	0.062 (0.241)	0.046 (0.209)	0.373
Poverty Rate[34]	15.29 (9.58)	14.90 (8.80)	15.14 (10.26)	0.791
Number of Cows	1.284 (1.239)	1.283 (1.135)	1.270 (1.396)	0.912
Number of Sheep	1.753 (3.477)	1.818 (3.311)	1.577 (2.366)	0.515
Number of Goats	1.438 (2.469)	1.039 (2.142)	1.717 (2.935)	0.019
Observations	1209	307	307	614

Table 3: Baseline household descriptive statistics, standard deviations are reported in parentheses

	Digelona Tijo	Doba	Goro Gutu	Gumay	Kersa	Meta	Tena	Overall Sample
Average Household Size	6.4	6.1	6.3	6.7	6.4	6.0	5.6	6.2
Average Total Flock Size	3.5	3.8	2.7	3.2	3.9	3.7	5.4	3.6

Table 4: Household Size and Flock Size by Woreda, Among Treatment and Control Households (n = 614)

[33] Reported p-values are calculated with standard errors clustered at the kebele level.
[34] Poverty rates are calculated at baseline using a poverty threshold of $1.00 per day at 2005 PPP.

As a result of EthioChicken's natural expansion plans, some woredas predominantly contained control households (such as Goro Gutu) while others mostly contained treatment households (such as Gumay, Tena, and Digelona Tijo). Table 5, below, shows the percentage of the treatment and control households in each woreda. The spatial distribution of the entire treatment and control sample can be seen in Figure 5.[35]

	Digelona Tijo	Doba	Goro Gutu	Gumay	Kersa	Meta	Tena	Total
Treatment	16.6%	20.5%	1.0%	17.9%	16.6%	20.5%	6.8%	100%
Control	9.1%	17.9%	26.7%	2.3%	25.1%	17.6%	1.3%	100%

Table 5: Distribution of treatment and control households, by woreda (n = 614)

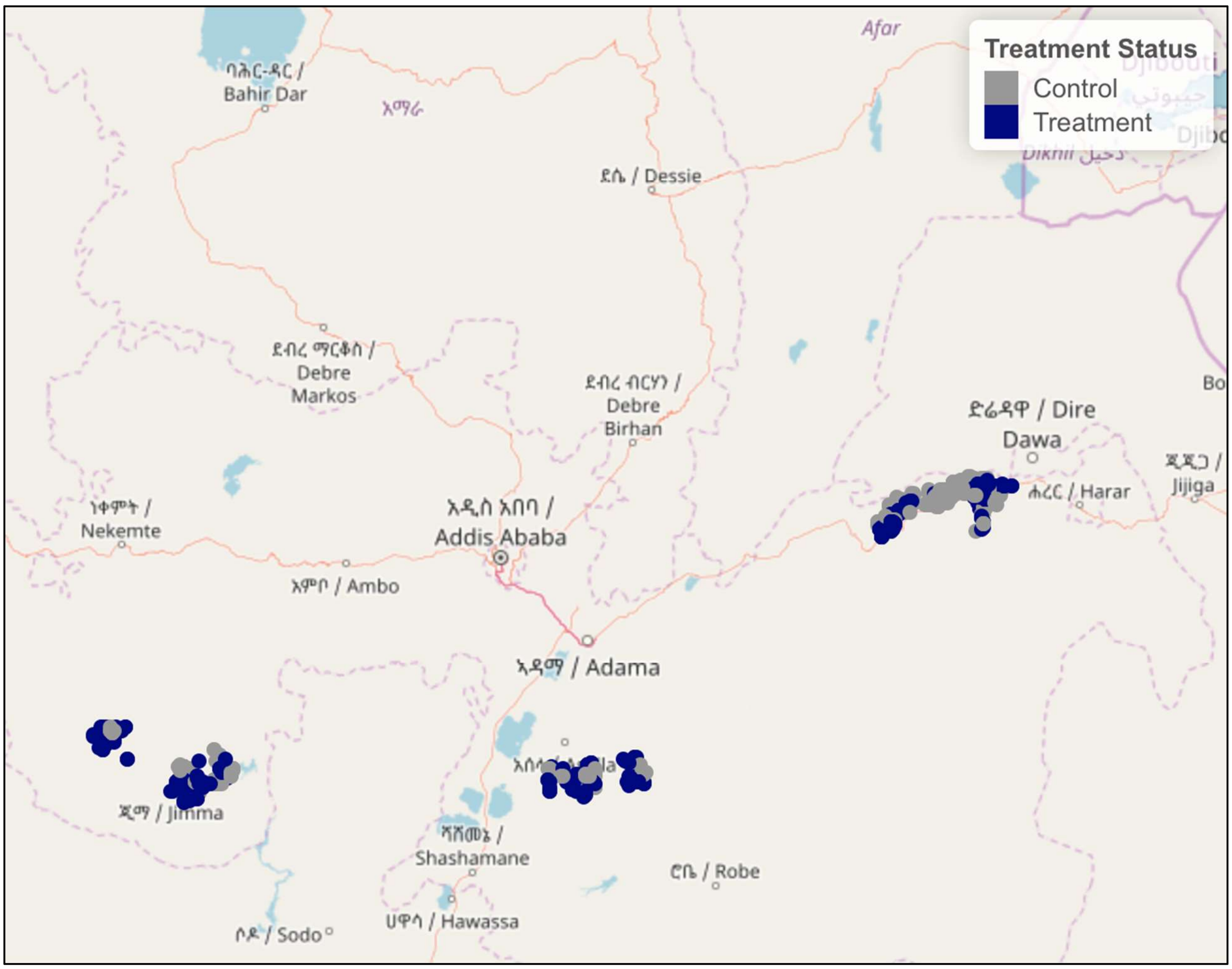

Figure 5: Map of treatment and control households[36]

[35] Due to uneven distribution of treatment and control by Woreda, we do not include woreda-level fixed effects in our analysis, as this would eliminate a large proportion of the variation in our sample. However, the main results are robust to including these fixed effects.

[36] An interactive version of this map can be downloaded here.

Across the entire sample of treatment and control households, 78% of households were Muslim, 20% were Orthodox Christian, and the remaining 2% were either other Christian denominations or mixed households. 96% of households in the treatment and control groups report growing crops, with maize, sorghum, and khat being identified as the most important crops.

Analytical Approach

To estimate the direct effect of chicken purchases, we use a standard ordinary least squares specification. We regress the outcome variable on an indicator for treatment, controlling for all of the variables we used for matching as well as for the baseline value of the outcome, if available. We also control for a dummy indicating whether the household was surveyed before or after Christmas. All errors are clustered at the kebele level.

In our analysis we include all households, even though some treatment households do not have ferenj chickens at the time of endline. This is to give a more holistic view of effects at the one year mark, taking into account that some outcomes (such as income from chicken sales) are likely to be felt even among household who no longer have chickens. In particular, it Is important to capture households that no longer own ferenj chickens in our sample because income from the sale of chickens is a vital component of overall revenue from chicken sources. In the Technical Appendix, we present results on our primary outcomes while restricting only to the sample of households that currently own ferenj chickens (and their matches). While some treatment effects are stronger in this sample, the overall pattern is quite similar to the results of the whole sample.

We measure the effect of ferenj chicken ownership on a number of outcomes guided by the theory of change. We first look at egg production and then income from chicken and eggs. Then we turn to nutrition, looking at food expenditure and consumption of eggs, meat, and protein for the household as well as for women and children specifically. Finally, we look at proxies of women's empowerment.

One drawback of our approach is that we did not attempt to measure total household income as this tends to be noisy and inaccurate for rural households. Therefore we are unable to say whether chicken-raising improves overall household income. Instead, we make the assumption that since chicken-raising take little time and money, any changes in chicken income and expenditure are likely additional and not substituting from other potential investments. Verifying this assumption would be an interesting goal of future research.

Additional details on the econometric specification can be found in the Technical Appendix.

Results

The first section presents a number of descriptive statistics of our sample while the second section presents impact estimates of key outcomes.

Descriptive Statistics

Chicken Ownership

Only treatment households own ferenj chickens, while both treatment and control households have habesha chickens. Among treatment households, 64.8% owned ferenj chickens at endline (as seen in

Figure 6) while 71.3% owned ferenj chickens in the prior 6 months before endline. In our treatment group, 69% of households have habesha chickens, while 62% of our control households have habesha chickens. (This difference is not statistically significant.)

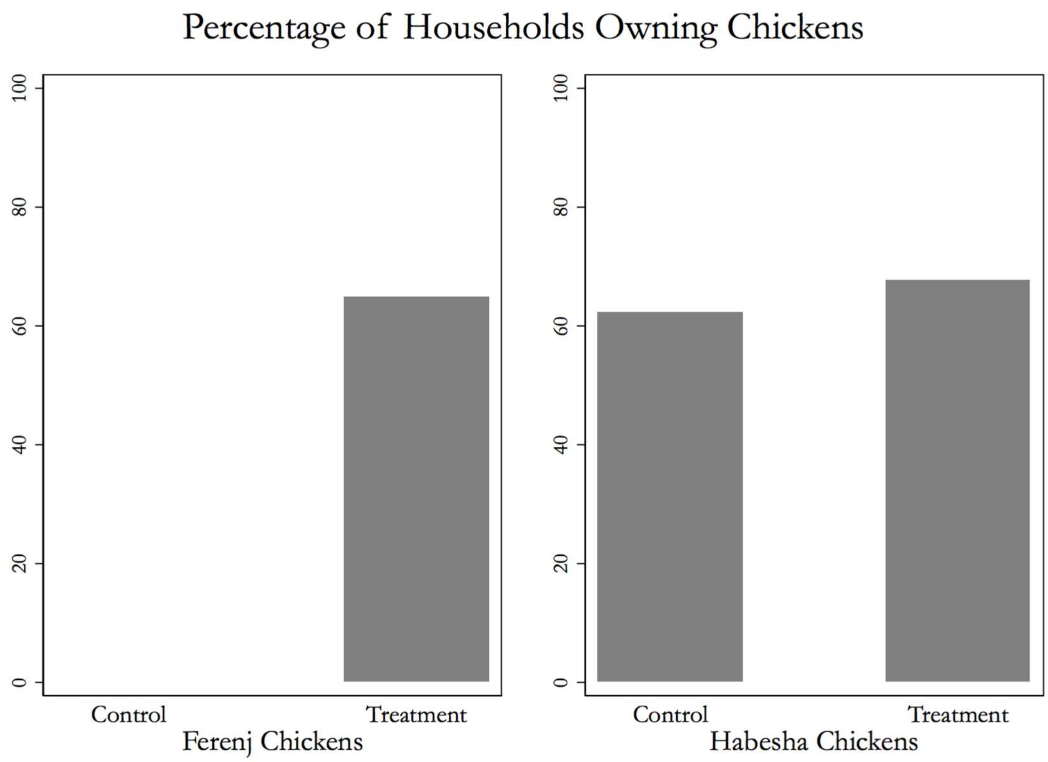

Figure 6: Chicken ownership by treatment status (n = 614)

Ferenj ownership seems to be *in addition to* habesha ownership rather than *in replacement of* habesha chickens. Treatment and control households own similar numbers of habesha chickens, with an average of 2.5 habesha chickens per household in each group. The median number of chickens owned by treatment households (4 chickens) is larger than for control households (2 chickens).

Figure 7 shows the distribution of the number of ferenj chickens owned by treatment households and the distribution of total chickens owned by households at the time of endline data collection (if a treatment household purchased ferenj chickens but sold them before endline, they would show up as owning zero ferenj chickens).

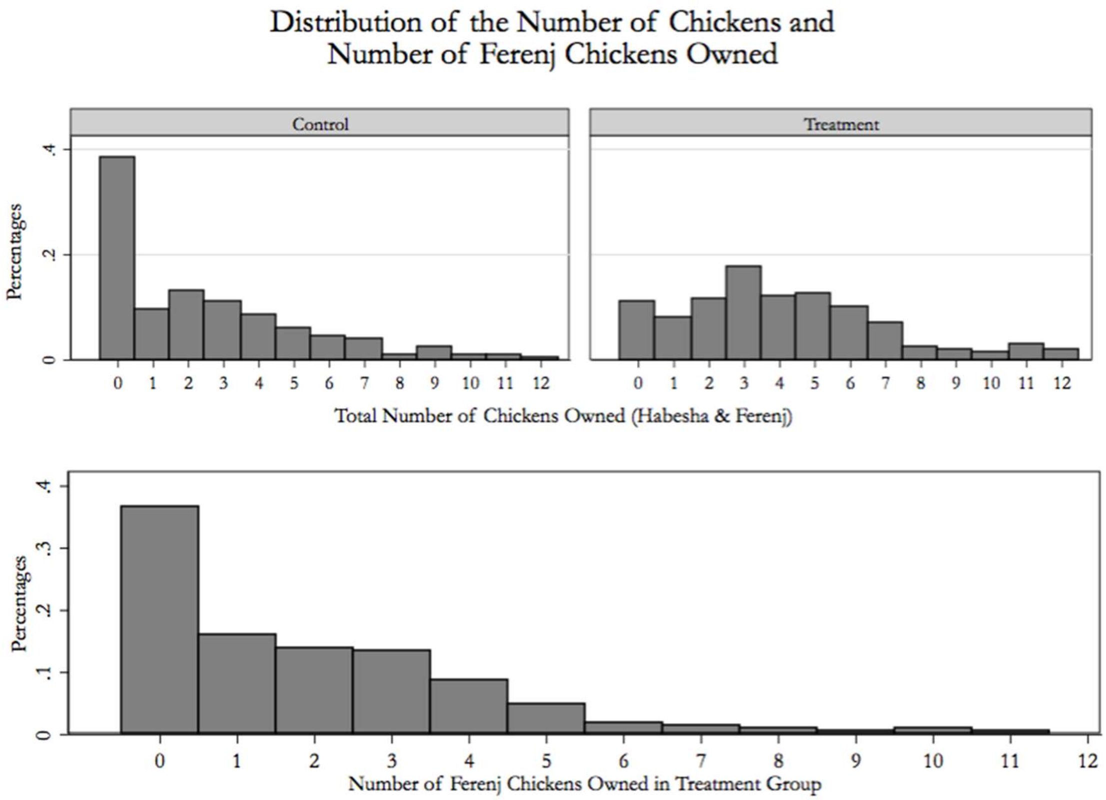

Figure 7: Distribution of chicken flock sizes among treatment and control groups (n = 307 in each panel). Figure truncated at N=12.

Table 6 shows the distribution of chicken breed and gender among treatment and control households.

	Female Habesha	Male Habesha	Bovans	Female Sassos	Male Sassos	Sasso, Unknown Sex	Unknown Ferenj Breed
Treatment Households (n = 307)	41.2%	12.4%	8.1%	12.4%	8.0%	0.3%	17.7%
Control Households (n = 307)	73.0%	27.0%	-	-	-	-	-

Table 6: Distribution of all chickens in sample, by chicken type and treatment status

Although enumerators were given photo cards to help them and the respondent identify the breed of ferenj chickens, this identification was certainly imperfect. Therefore, it is likely that some of the “unknown” breeds of ferenj chickens were actually Sasso or Bovans, and potentially were also from EthioChicken.

Table 7 presents summary statistics on breed performance, confirming that ferenj chickens are more productive. As expected, ferenj chickens produced more eggs per week on average [ferenj chickens produced an average of 5 eggs per week while habesha chickens produced an average of 2.6 eggs per week]. This is partially due to the fact that ferenj chickens lay eggs more consistently. 65.3 percent of habesha owners reported that their chickens had stopped laying eggs for a period of more than 5 days after having reached egg laying age while only 31.9 percent of ferenj owners reported the same. The statistics in this table are presented for the 849 households that comprise the direct effect and spillover samples. The figures on egg production are a blended average of times when chickens are laying and times when they are unproductive.[37]

Statistic	Habesha	Ferenj
Average eggs produced per week	2.63	5.04
Median egg price	3 ETB	3 ETB
Percentage of households that lost a chicken of this type to disease	43.7	37.3
Percentage of households whose chickens of this type stopped laying eggs for more than 5 days	65.3	31.9
Average price received for selling a chicken of this type	103.2 ETB	189.0 ETB

Table 7: Habesha and ferenj performance (n = 849)

There is no difference between the median egg price of habesha and ferenj eggs (the median price for both habesha and ferenj eggs is 3 ETB), but there is a difference in price between habesha and ferenj chickens sold at market.. Ferenj chickens receive a greater price when sold at market, likely due to their larger size. On average, habesha chickens received a price of 103.2 ETB while ferenj chickens received a price of 189.0 ETB.

There is a moderate difference in reported mortality: among households that owned habesha chickens within the last six months, 43.7 percent reported losing at least one habesha chicken to disease, while 37.3 percent of households that owned ferenj chickens reported at least one ferenj chicken dying because of disease (a difference that is significant at the 0.10 level).

Table 8 presents information on egg production and egg uses for treatment and control households. Importantly, treatment households have an average of 3.1 additional missing eggs per week than control households.[38] We define “missing eggs” as the difference between reported egg production and reported egg uses on our survey. We believe this difference in missing eggs likely results from two sources: (1) heightened recall bias among treatment households, and (2) other egg uses that weren’t captured in the survey. On the first point, treatment households owned more chickens and ferenj chickens are higher productivity. Since treatment households have more eggs, respondents in the treatment group may have greater difficulty remember exact production figures and how eggs were used. To investigate other egg

---

[37] Rather than discriminating between chickens that were laying and those that had stopped laying at the time of the endline survey, households were asked “How many eggs did your –FERENJ or HABESHA- chicken produce, in total, in the last 7 days (one week)?”

[38] The presence of these missing eggs in the treatment group may lead us to underestimate the effects of treatment on egg sale or egg consumption, if some of the missing eggs actually belong in those categories. There is no principled justification to attribute these missing eggs to one usage category or another, so our overall results on consumption and sales do not incorporate any fraction of the missing eggs in their estimates.

uses that weren't captured in the survey, we asked Acumen's Lean Data team to contact ferenj owners to explore other eggs uses beyond consumption, sales, given away, and hatching. Some households reported that these other eggs broke, were consumed by animals, or were used for bartering. However, the amount of eggs that fell into these other categories tended to be small relative to the main categories of consumption and sales.

	Total Eggs Produced	Total Eggs Consumed	Total Eggs Sold	Total Eggs Given Away	Total Eggs Kept for Hatching	Total Eggs Missing
Treatment Households (n = 307)	11.3	6.6	1.0	0.2	0.7	2.9
Control Households (n = 307)	4.5	3.6	0.6	0.1	0.4	-0.2

Table 8: Weekly Egg Production and Uses for Treatment and Control Households

In Figure 8, we show households' responses to general questions about their perceptions and attitudes towards ferenj chickens. Most households in the sample expressed positive sentiments about raising ferenj chickens, with the majority expressing strong agreement that ferenj chickens are a good investment and that they prefer raising ferenj chickens over habesha chickens. However, access to ferenj chickens was more of an issue. Even in the treatment group, over 60% of households disagreed with the statement "The process to purchase a ferenj chicken is easy." However, treatment households were more likely than control households to know how to purchase a ferenj chicken and believe it is a good investment. While the difference between treatment and control is most pronounced for these two categories, treatment households also have slightly better perceptions of ferenj chickens on the other four attitudinal dimensions reported below.

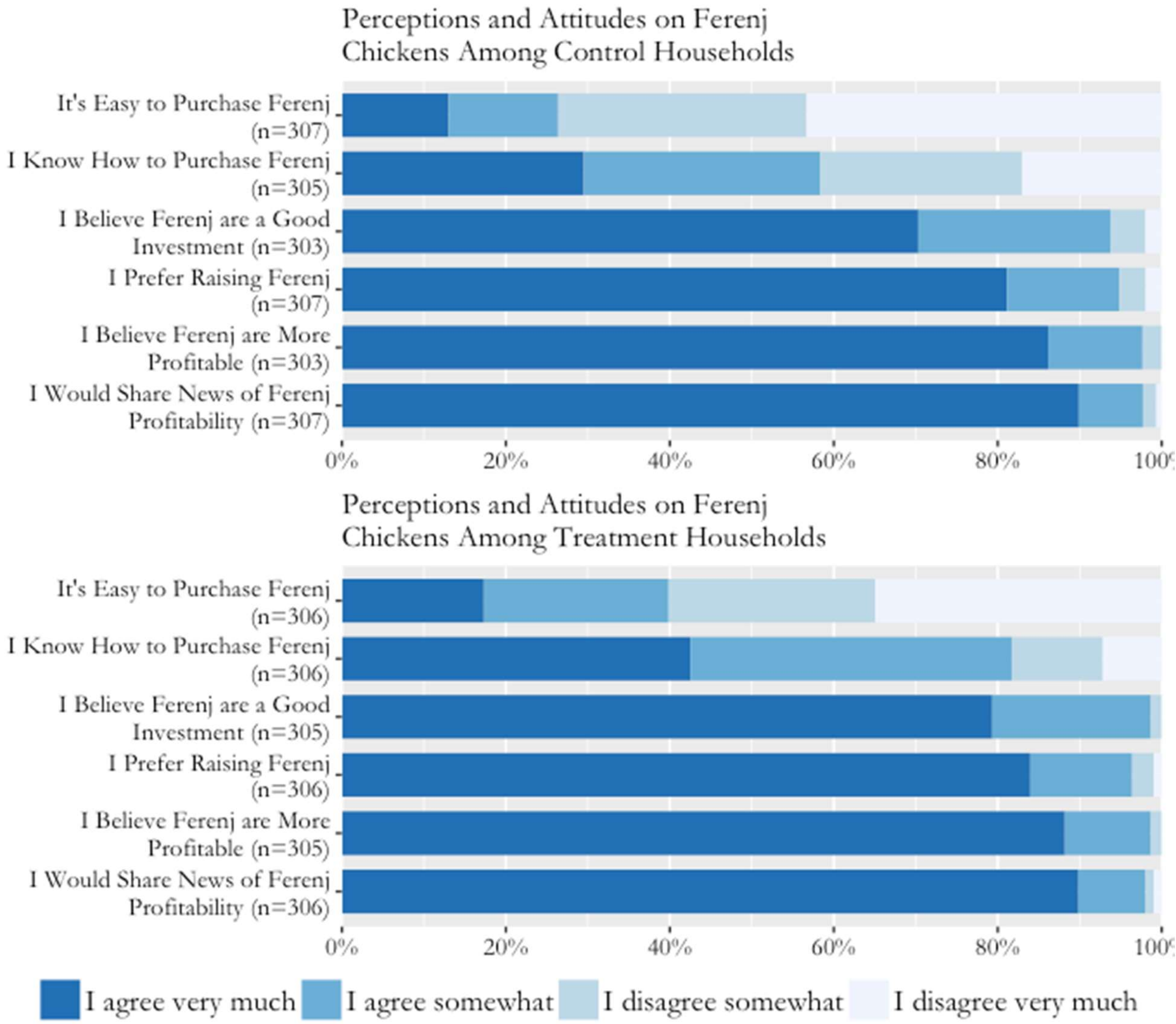

Figure 8: Perceptions and Attitudes on Ferenj Chickens

Figure 9 shows other preferences around purchasing chickens. Households in the direct and spillover samples are most likely to prefer purchasing chickens during the Bega (Winter) season. However, contrary to popular belief, households do not report a particular reluctance to purchase chickens during the Kiremt (Summer, rainy) season. Approximately 50% of households say they would be happy to purchase chickens during the Kirmet season and over 30% report that they have purchased during Kirmet before.

As shown in Figure 10, Development Agents (DAs) were the most important influencers of the decision to purchase ferenj chickens, with about 20% of households saying their decision to raise chickens had been influenced by their local DA. Family, friends, and other household members also exerted some influence over this decision while neighbors, other local government officials, and other individuals influenced the decision to a lesser extent. Over 60% of households report that nobody in particular influenced their decision.

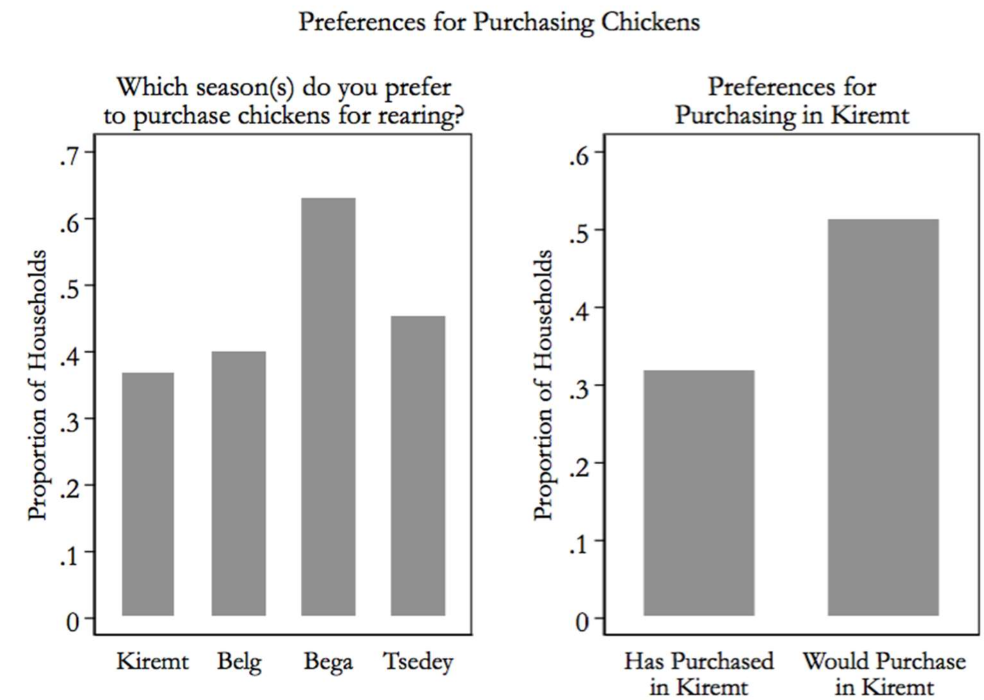

Figure 9: Preferences for Purchasing Chickens among direct and spillover samples (n = 849)

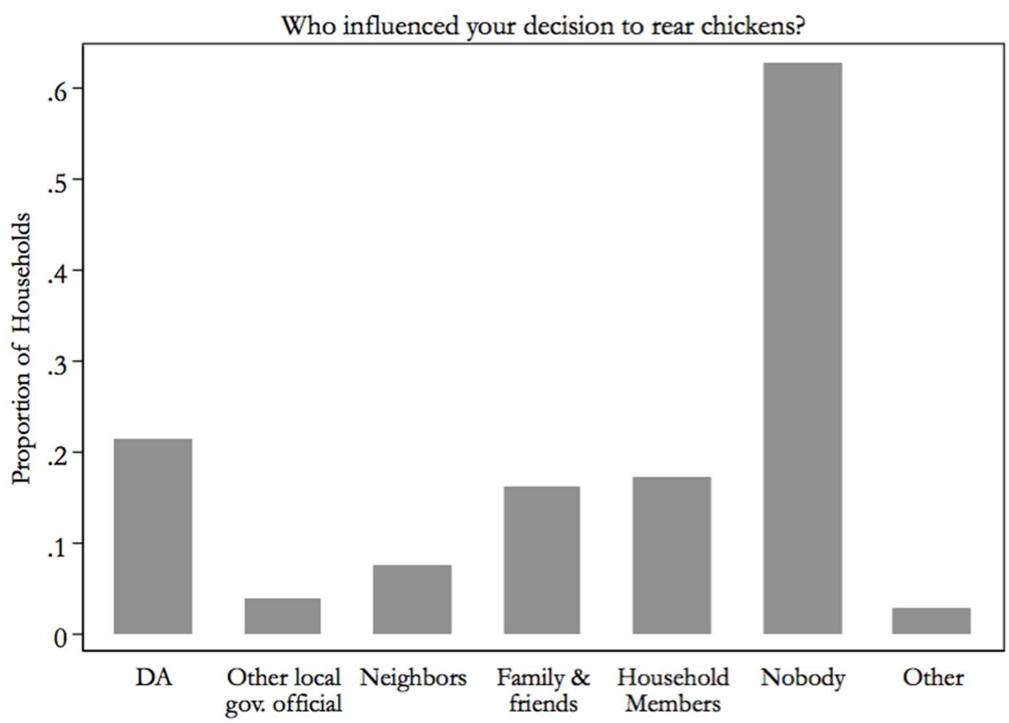

Figure 10: Influencers of decision to purchase ferenj chickens among direct and spillover samples (n = 849)

Among the combined direct effect and spillover sample, 105 households owned a ferenj chicken at one point but stopped owning before the endline. Figure 11 graphs various reasons given for stopping ferenj rearing. The most common response, from 70 percent of households that no longer had a ferenj chicken, was that their chicken(s) had died.

Household were also asked what factors might induce them to eat more ferenj chickens. 75 percent of households in the combined sample said they would eat more ferenj chickens if they owned more ferenj chickens. As seen in Figure 12, households also cited lack of money, lack of ferenj availability, and the cost of ferenj chickens as important barriers limiting their ferenj consumption.

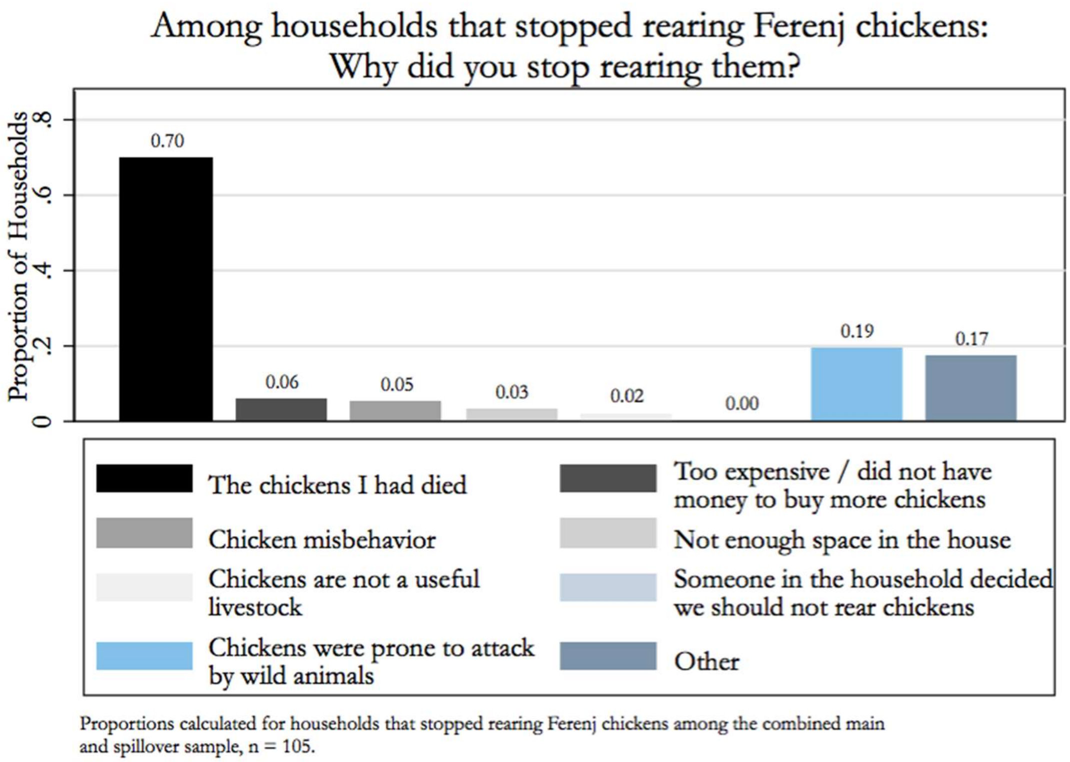

Figure 11: Reasons households stopped owning ferenj chickens among direct and spillover samples (n = 849)

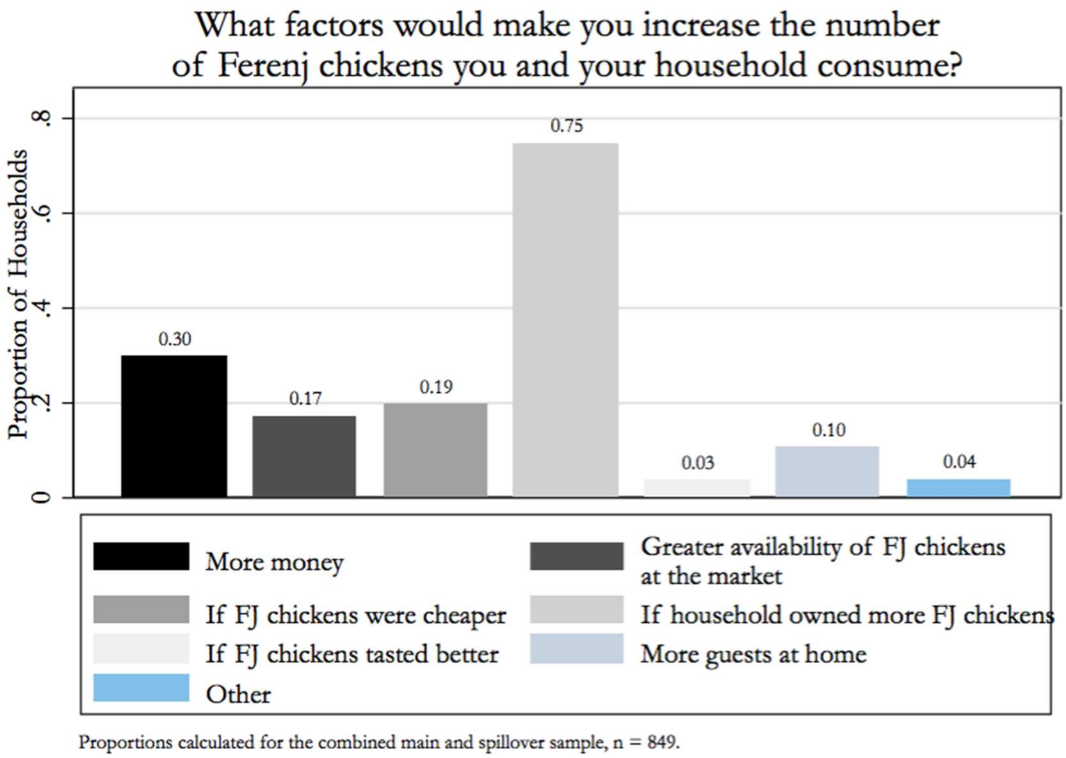

Figure 12: Factors that would increase ferenj consumption among direct and spillover samples (n = 849)

Impact Estimates

We now turn to estimating the impact of ferenj purchases on our outcomes of interest.

Impact results are presented as figures, which show the average outcomes of the treatment and control groups. The difference in the height of the bars represents the estimated effect of ferenj chicken ownership. The black line on the treatment bar represents the 90% confidence interval of the treatment mean, which indicates a reasonable margin of error of the impact estimate. No intersection between this bar and the control mean signifies a statistically significant treatment effect, denoting a high likelihood that treatment and control group results for that outcome are different. Additionally, the number of stars next to the treatment mean in the graphs designates the statistical significance of the result.

All averages and impact estimates presented in these results account for differences in covariates (explained in detail in the regression analysis section). As such, the averages presented here may vary slightly from actual averages of outcomes that do not adjust for differences in covariates. However, a simple comparison of means yields similar results in all cases.

Detailed regression tables can be found in the Technical Appendix.

Egg Production

We first look at the impact on egg production. We find that treatment households increase egg production to 11.3 eggs per week, compared to production of 4.5 eggs/week in the control group.

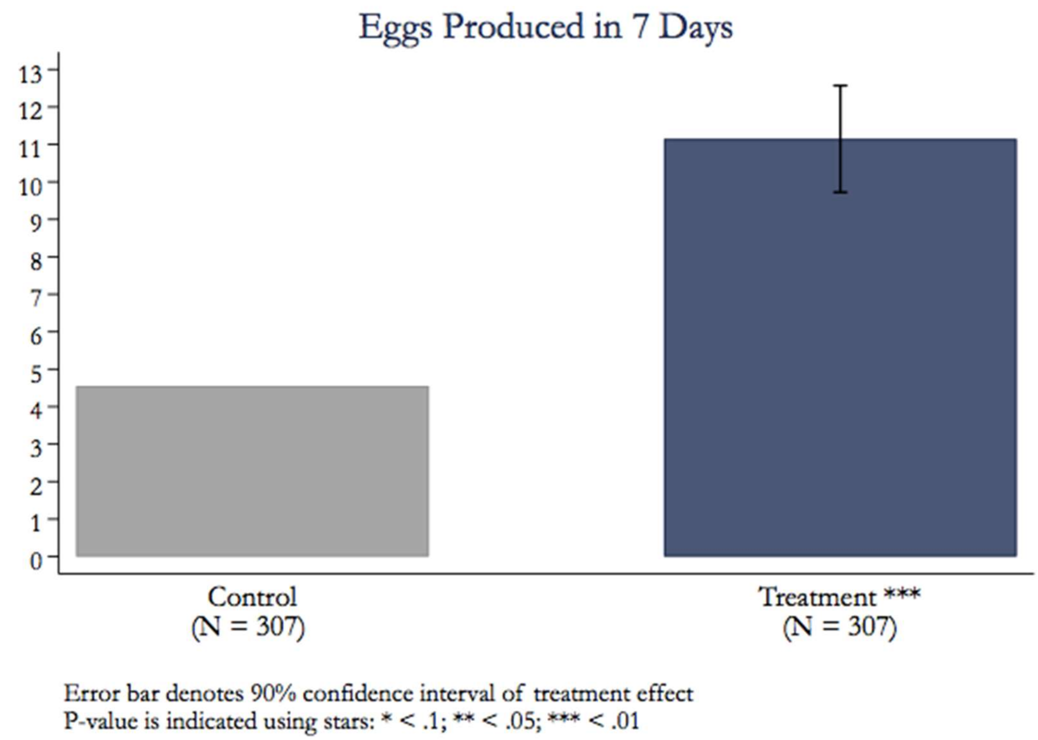

Figure 13: Impact estimate - eggs produced in last 7 days

This treatment effect (of 6.6 eggs/week) is due both to treatment households having larger total flock sizes, as well as ferenj chickens being more productive. We find that treatment households have around the same amount of habesha chickens, but add around 2 ferenj chickens to their flock. Households with an additional habesha chicken have on average an additional two eggs per week. Therefore, had the households increased their flock size by two habesha chickens we would have expected an increase of around 4 eggs per week. The additional 2.8 eggs per week comes from the fact that ferenj chickens are more productive.[39] As would be expected, the treatment effect on egg production increases with the size of the ferenj flock (see subgroup analysis in the appendix).

Income from Chicken and Eggs

Households can generate income from poultry in two ways: selling eggs and selling chickens. Figure 14 shows that treatment households do indeed sell more eggs, reporting an increase of 4.5 eggs sold in the

[39] While the above decomposition comes from a loose calculation, a more formal structural model (implemented with STATA's 'sem' command) gives a very similar result. It finds that the increase in flock size contributes to an increase of approximately 4.5 eggs, while increased ferenj productivity contributes to an increase of an addition 2.24 eggs per week.

previous month, compared to 2.5 eggs in the control group. This results in an increase in income of 14 ETB per month (approximately $0.50) over a control-group average of 7.2 ETB ($0.26).[40]

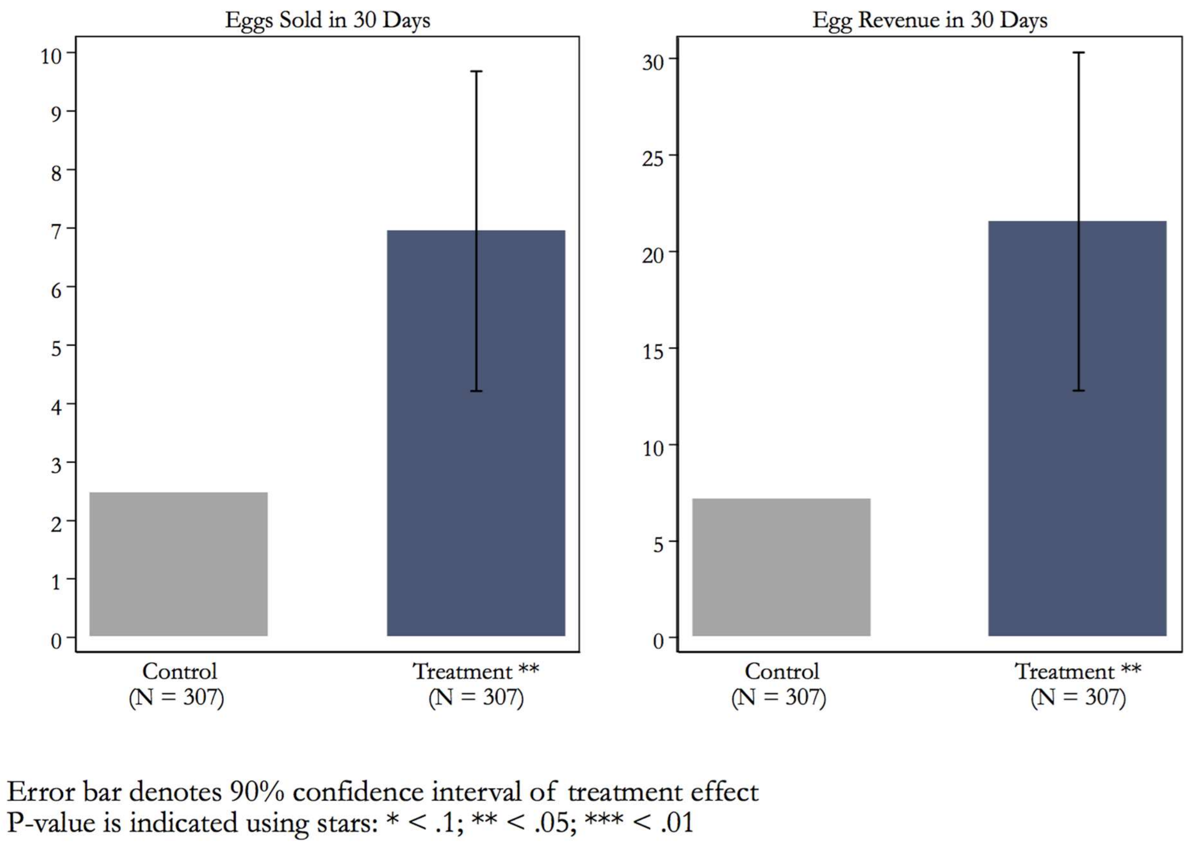

Figure 14: Impact estimate - egg sales and egg revenue

Treatment households almost double their revenue from selling chickens—mainly due to the higher price of ferenj chickens at market—but the actual gains from sale are limited due to the fact that most households only sold 1-2 chickens. Figure 15 shows that treatment households sell more chickens (though this result is not statistically significant at conventional levels), resulting in higher revenue. As with egg production, a larger number of chicken sales may be a result of larger flock sizes in treatment households rather than a higher likelihood of selling ferenj chickens. Indeed, if we control for flock size, the effect of the treatment on chicken sales goes to zero. Increasing flock size by one chicken—whether that chicken is habesha or ferenj—is associated with an increase in the number of chickens sold by 0.20. Treatment households report increased revenue from chicken sales of 88 ETB (3.19 USD) over the course of the past 6 months. This is an increase over an average of around 103 ETB (3.74 USD) in the control group. Ferenj chickens are sold for an average of 189 ETB, compared to a purchase price of 74 ETB for Sassos and 78 ETB for Bovans. The impact on revenue from chicken sales is higher among Christian

[40] Note that besides selling eggs and consuming them (which is discussed later), households can also give eggs away or keep them for hatching. We find that treatment households keep one more egg for hatching compared to control, and give away an average of an additional half an egg per week.

households than in Muslim households, though this difference may reflect variation in the timing of access to ferenj chickens rather than a "religious effect."[41]

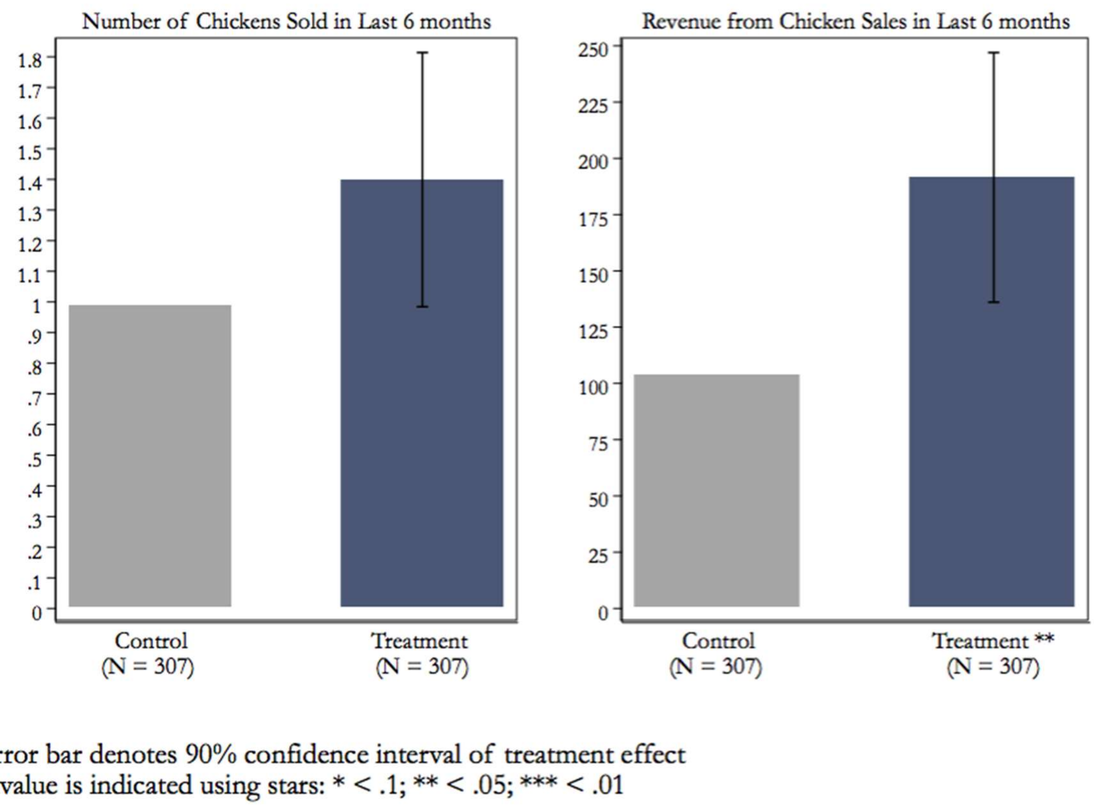

Figure 15: Impact estimate - chickens sold and chicken revenue

Chicken Expenses

We construct estimates of a household's total expenses related to chickens for the month preceding endline data collection. As seen in Figure 16, treatment households reported higher average total expenses of 35.5 ETB compared to 7.2 ETB among control households. That being said, median expenses in the treatment group (and the sample as a whole) were less than 2 ETB per month. This suggests that most households in both treatment and control groups spend minimal money on raising their chickens, but some households incur significant expenses.

[41] Within our sample, Muslim households tend to be clustered around East and West Hararghe—if these areas received ferenj chickens later than the other areas in the study, we might be picking up an effect caused by the timing of distribution rather than religion (i.e. areas that received the chickens earlier would have had more time to sell their chickens than households that only received the chickens recently).

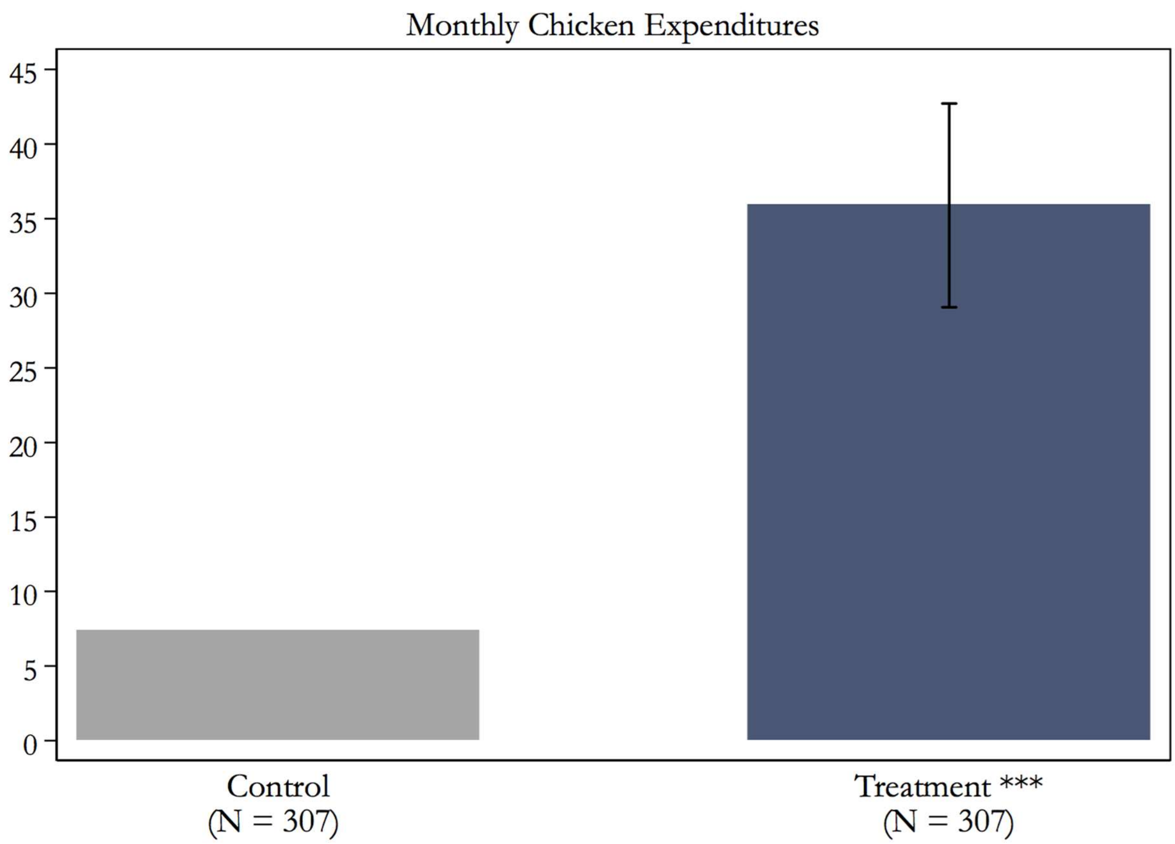

Figure 16: Impact estimate - monthly chicken expenses

To get a better picture of household expenses, Table 9 reports monthly expense levels at different points of the distribution for treatment and control households. Expenses in this table are reported *per chicken* rather than *in total* to isolate change in costs from the fact that treatment households have larger flocks. As this table shows, the difference in expenses between treatment and control households comes about only in the top half of the distribution; in that half of the distribution there are quite large differences between those groups. Both the increase in flock size and the top half of treatment households having meaningfully larger expenses per chicken drive the large treatment effect on total expenses observed in Figure 16 above.

	Expense Per Chicken, Control Households (ETB)	Expense Per Chicken, Treatment Households (ETB)
25th Percentile	0	0
50th Percentile	0	0.28
75th Percentile	0	11.0
95th Percentile	12.5	64.2

Table 9: Distribution of chicken expenses, by treatment status

Table 10 breaks down total chicken expenses per month into nine expense categories.[42] Looking at average expenses, feed is the single largest expense for both treatment and control households (at 47% of all expenses for control households and 45% of all expenses for treatment households) and treatment households pay over double per chicken on feed than control households. Treatment households spend slightly more on all remaining categories except 'other expenses', but increased expenditures on transportation for selling eggs, antibiotics, and veterinary expenses are of particular note. Higher spending on transportation to sell eggs might be expected among the treatment group, though, given their higher overall egg productivity.

However, looking at median expenses paints a very different story—in fact, among both treatment and control households, the median expense for every single expense category is 0 ETB. This provides further evidence for most households having minimal, if any, expenses on chickens. The large effect on total monthly chicken expenses is driven entirely by a minority of households in the treatment group.

	Average Expense Per Chicken, Control Group	Median Expense Per Chicken, Control Group	Average Expense Per Chicken, Treatment Group	Median Expense Per Chicken, Treatment Group
Feed	4.0 ETB	0 ETB	9.4 ETB	0 ETB
Transportation for Purchasing Chickens or Eggs	0.2 ETB	0 ETB	0.4 ETB	0 ETB
Transportation for Selling Eggs	0.5 ETB	0 ETB	1.9 ETB	0 ETB
Transportation for Selling Chickens	0.1 ETB	0 ETB	0.1 ETB	0 ETB
Vaccinations	0.4 ETB	0 ETB	1.0 ETB	0 ETB
Antibiotics	2.0 ETB	0 ETB	5.4 ETB	0 ETB
Vitamins	0.3 ETB	0 ETB	0.9 ETB	0 ETB
Veterinary Expenses	0.1 ETB	0 ETB	1.6 ETB	0 ETB
Other Expenses	0.9 ETB	0 ETB	0.3 ETB	0 ETB
Total	8.5 ETB	0 ETB	21.0 ETB	0 ETB

Table 10: Breakdown of expenses, by treatment status

Since chicken feed is such a large component of total chicken expenses, Table 11 reports on the purchase of different types of feed among treatment and control households. Households could choose multiple responses for this question. In both groups, the majority of households did not purchase feed. However, a

[42] Since vaccinations, antibiotics, vitamins, veterinary expenses, and other expenses may be infrequent, they were asked over a recall period of 6 months in the survey but have been scaled to one month here. There is a tradeoff between the accuracy and precision of these measurements. Some activities, such as vaccination or visiting a veterinarian, are likely to occur infrequently. Asking for certain expenses over a recall period of 6 months increases the likelihood that we will capture legitimate expenses that did not occur recently yet – due to the potential elapsed time between that expenditure and endline data collection – the reported figure may not be very precise. This tradeoff applies to smaller recall periods as well: we asked about feed purchases in the previous month, but if the only food expenditures occurred at the beginning of that month, the resulting measures may lack precision. However, there is no reason we would expect any potential recall bias to systematically affect our measurement in either direction.

larger percentage of treatment households purchased feed of each type. Therefore, while most households in both groups have zero expenses on feed, treatment households report higher average feed costs per chicken as a result of the remaining households being more likely to purchase feed.

	Percentage of Control Households	Percentage of Treatment Households
Does Not Purchase Feed	94.1 %	73.6 %
Grains	5.2 %	18.6 %
Corn / Maize	1.0 %	5.2 %
Frushka (leftovers from the mill)	0.0 %	4.9 %
Store bought, packaged feed	0.0 %	2.6 %
Mixed feed, unpackaged	0.0 %	0.7 %

Table 11: Types of feed used by households

These results indicate there is wide variation in chicken expenses, particularly among the treatment group. In light of this, the average treatment effect seen in Figure 16 likely obscures significant heterogeneity. Table 12 reports quantile regression output for total monthly expenses (not per chicken expenses) at various points of the distribution. For most households, purchasing ferenj chickens has zero effect or minimal effects on monthly chicken expenses; however, there is a larger effect on households in the top of the expense distribution. The average treatment effect, included in the last column of the table as a point of comparison, does not reflect the reality for the majority of households.

Point in the Distribution of Monthly Expenses	20th Percentile	40th Percentile	60th Percentile	80th Percentile	Mean (OLS)
Magnitude of Treatment Effect	0.0	0.0	5.3	48.6	28.5

Table 12: Quantile regression estimates - chicken expenses

Chicken Profit

Chicken revenue can vary over time depending on when households sell chickens and therefore it can be difficult to get an accurate measure of profit on chicken-raising at any given point in time. We estimate profit as the difference between revenue and expenses in the month preceding endline data collection, which were discussed in the prior section. This analysis therefore provides a snapshot of profits, but may not reflect chicken profits over the lifecycle of a flock. Revenue includes income from eggs and chicken sales, with chicken sales as the number of chickens sold in the last 6 months scaled to a single month by dividing by six.

Although revenues related to chicken increased, treatment households actually have lower average monthly profit than control households as a result of increased chicken expenses. As shown in Figure 17, treatment households have lower profits by around 11 ETB/month. The magnitude of this effect increases with ferenj flock size: as seen in subgroup analyses of the appendix, households with more than 4 or more than 5 ferenj chickens have a larger decrease in profit compared to the overall sample of households. The subgroup analyses also show that the treatment effect on profitability is more negative among the subset of households that owned ferenj chickens at endline data collection or within six months of endline.

However, there are a number of reasons to interpret the observed effect on profit with caution. First, by design our treatment group is made up of many people who had purchased their first batch of ferenj chickens in the previous year. Therefore many of them likely made investments in these chickens but have not yet reaped the benefits they will get when they eventually sell them. This is likely not true in control households. This imbalance can be seen in the duration of chicken ownership: 74% of households in the direct and spillover sample who owned habesha chickens indicated they had owned at least one habesha chicken for greater than one year while only 35% of ferenj-owning households said the same about their ferenj chickens. This suggests that many ferenj owners might not have owned their ferenj chickens for long enough to reap their full rewards. Additionally, these profit numbers do not take into account consumption of home-produced poultry products. If treatment households are saving money on egg expenditures by producing their own eggs (which seems borne out in the expenditure data), then this is a monetary benefit not captured in construction of the profit variable.

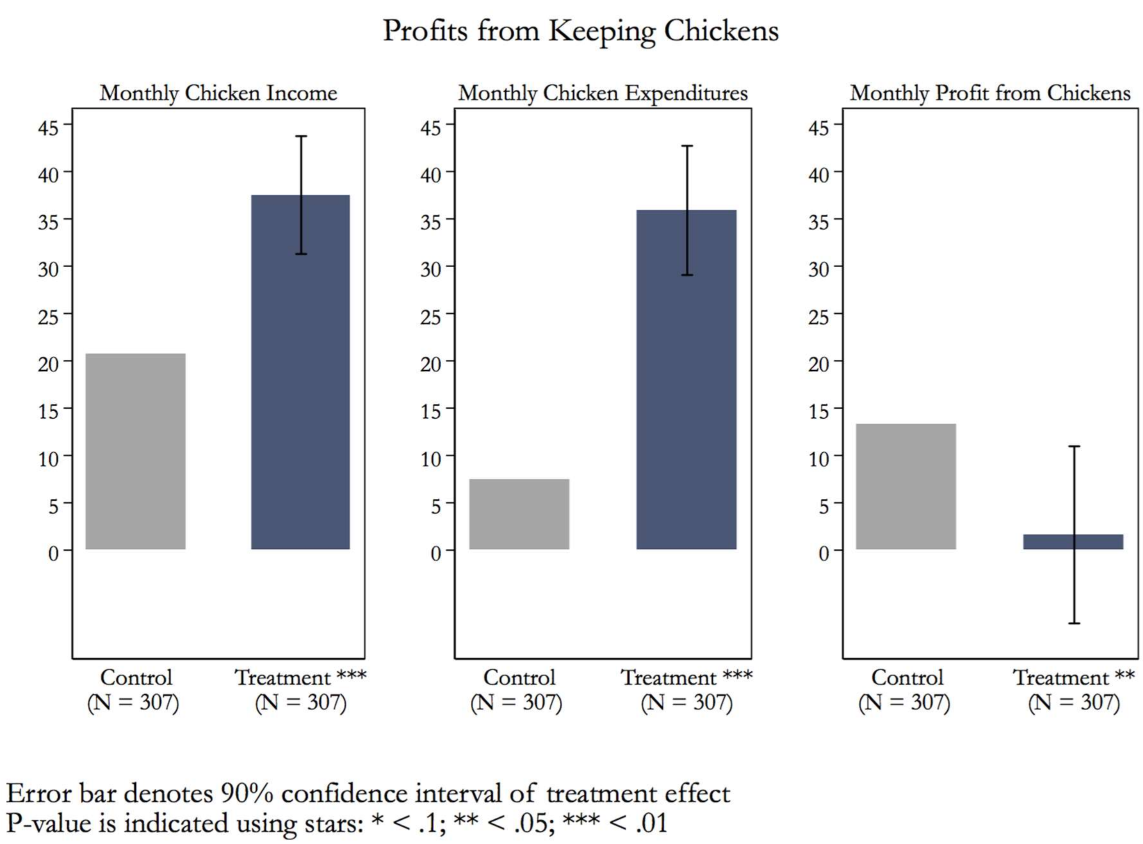

Figure 17: Impact estimate - chicken expenses and profits

As with expenditures, the average treatment effect may mask important variation within the distribution of chicken profits. Table 13 reports quantile regression results at various points of the chicken profit distribution. The majority of households in the sample are not made any better off by the treatment, though some households do reap benefits. The bottom end of the distribution performs poorly, the middle of the distribution sees no effect on profits, and the top end of the distribution sees marginal gains.

Point in the Distribution of Monthly Profits	20th Percentile	40th Percentile	60th Percentile	80th Percentile	Mean (OLS)
Magnitude of Treatment Effect	-27.2	-0.5	0.0	13.5	-11.6

Table 13: Quantile regression estimates - profits from rearing chickens

Table 14, below, provides information on the profiles of profitable and unprofitable households. Profitable households are defined as those with profits greater than zero while unprofitable households are defined as those with profits less than zero. On average, profitable households in the sample of treatment and control households had slightly larger flocks, slightly more habesha chickens, greater income, fewer expenses, larger egg sales, and larger chicken sales (results that are significant at the 0.01 level). Selling both chickens and eggs seems to be a crucial driver for the overall profitability of households. There is no evidence of a difference in ferenj flock size between profitable and unprofitable households.

	Total Flock Size	Ferenj Flock Size	Habesha Flock Size	Chicken Income (monthly)	Chicken Expenses (monthly)	Eggs Sold (30 days)	Chickens Sold (6 months)
Profitable Households (n=188)	4.42	0.97	3.45	70.95	7.36	8.70	3.11
Unprofitable Households (n=426)	3.26	1.15	2.11	7.77	27.69	1.13	0.41
Difference	1.16***	-0.18	1.34***	63.18***	-20.33***	7.58***	2.69***
*** indicates that results from a two-tailed t-test are significant at the 0.01 level.							

Table 14: Profile of Profitable and Unprofitable Households among Treatment and Control Groups

As mentioned above, the negative treatment effect on profits may also partly be a consequence of ferenj-owning households not having owned chickens long enough for them to sell their higher-productivity chickens. Figure 18 shows an analysis where the treatment group is limited to treatment households that have actually sold at least one ferenj chicken and control households that have sold at least one habesha chicken. The results from this subgroup investigation are more promising than those depicted in Figure 17: although chicken expenditures among the treatment group are still greater than among the control group, there is a positive treatment effect on monthly profits. Average profits (not covariate-adjusted) among the treatment subgroup are 59.0 ETB per month while average profits among the control group remain at an average of 48.4 ETB per month. Although the data does provide some insight into the relationship between chicken ownership cycle and profits, this is suggestive evidence that a full understanding of the profitability of ferenj chickens requires a longer evaluation timeline. Although these restricted results are more positive, the practical significance of the treatment effect is still limited—with a covariate-adjusted difference between treatment and control of 19.4 ETB (0.71 USD) per month, the household would only be able to afford one additional Sasso chick from EthioChicken approximately once every 3.8 months.

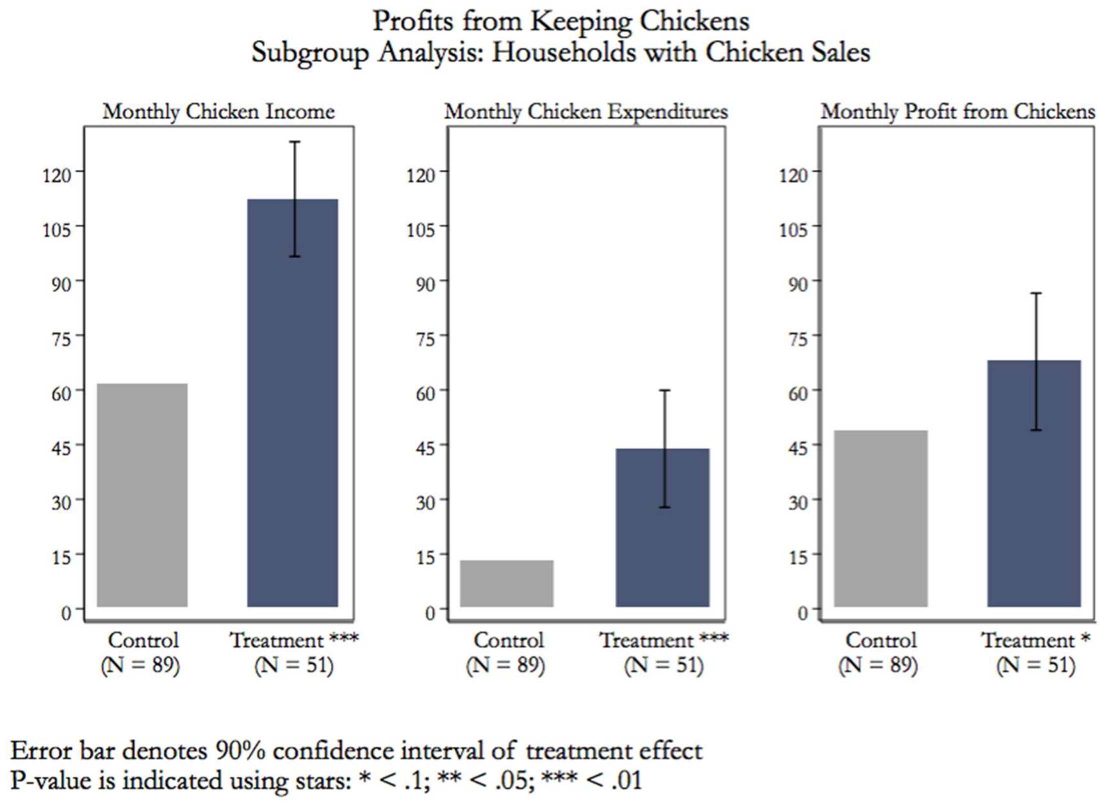

Figure 18: Impact estimate - chicken expenses and profits for subgroup with chicken sales

Modeling Per-Chicken Profitability

To develop a broader picture of chicken profitability, Table 15 presents a basic profit model for each of the main chicken breeds. As we do not have data on duration of chicken ownership, this model assesses profits for female chickens at the end of one year of ownership since the female chicken began laying eggs.[43] We did not disaggregate data on egg production or sales price for different breeds of female ferenj chickens, so our profit model includes an aggregated entry for "Female Ferenj". For male ferenj chickens, we assess income, expenses, and profits at the point when they reach market weight and could be sold for meat—approximately 3 months after purchase.[44] Male habesha chickens take approximately one year to reach market weight.[45] We also include a row for "Male Sasso (yearly)" which assesses profitability under the assumption that households are able to cycle through 4 purchases and sales of male Sasso chickens in one year. Although this represents a theoretical maximum profit for chicken-

[43] Due to the recent distribution of ferenj chickens in these areas, we do not have information on duration of ownership. However, 74% of habesha owners in the direct and spillover samples indicated they had owned at least one of their chickens for at greater than one year. On this basis, assessing profits for female chickens at the year mark seems reasonable.

[44] Aglionby, John. "EthioChicken: Ethiopia's well-hatched idea." *Financial Times*. Nikkei Inc., 16 March 2018.

[45] Ibid.

rearing, it should be estimated with caution given that Acumen's Lean Data studies showed that households tend to purchase chickens only once a year. Table 15 includes separate rows for average and median statistics, which are calculated for households in the combined direct and spillover samples.

Chicken Type	Statistic	Cost of Chick (ETB)	Income from egg sales (ETB)	Income from selling chicken (ETB)	Feed Expenses (ETB)	All Other Expenses (ETB)	Profit After Sale (ETB)
Male Habesha	Average	0	0	87.7	42.4	13.9	31.4
	Median	0	0	85	0	0	85
Female Habesha	Average	0	29.2	87.7	42.4	13.9	60.6
	Median	0	0	85	0	0	85
Female Ferenj	Average	76	64.3	160.7	118.8	45.1	-14.9
	Median	76	0	153	0	0	77
Male Sasso	Average	74	0	160.7	29.7	11.3	45.7
	Median	74	0	153	0	0	79
Male Sasso (yearly)	Average	296	0	642.8	118.8	45.1	182.9
	Median	296	0	612	0	0	316

Table 15: Per-chicken profit model for habesha and ferenj breeds (using data from direct effect and spillover samples)

The profit models for these breeds are meant to be illustrative rather than definitive as they rely on several assumptions (included in the footnotes).[46] However, the pattern observed here mirrors the treatment effect on profit observed in Figure 17: when looking at sample averages, this analysis suggests that owning ferenj chickens is unlikely to be more profitable than owning habesha chickens (given that females composed 76 percent of household ferenj flocks). On a per-chicken basis, male ferenj chickens are slightly more profitable than male habesha chickens but female ferenj chickens are considerably less profitable than their habesha counterparts. Profits from owning and selling four male Sassos over the course of one year are much higher than those from owning and selling one male habesha.

As previously discussed, median values for income and expenses are much different from average values. The results when using median values are more positive for ferenj breeds in terms of absolute profits but individual male and female ferenj chickens are less profitable than their habesha counterparts . Median

[46] One assumption we include in the model is a mortality rate between the acquisition of chickens and when those chickens reach a productive age of 15 percent. The mortality rate is approximately 16% for Sassos and approximately 14% for Bovans. We simplify this to an overall 15% mortality rate for ferenj breeds. To incorporate this information into the expected profitability of a chicken, we assume the mortality rates for habesha and ferenj breeds are the same and scale all columns by 0.85 except the purchase cost for a chick. Getiso, A. et al. "Production performance of Sasso (distributed by EthioChicken private poultry farms) and Bovans brown chickens breed under village production system in three agro-ecologies of Southern Nations, Nationalities and Peoples' Regional State (SNNPR), Ethiopia." *International Journal of Livestock Production.* 8.9 (2017) 145-157.

egg income and median expenses for all breed types are zero. Though ferenj breeds are sold for more money at market, the greater start-up cost of purchasing ferenj chickens pushes their profitability below the habesha breeds (except when looking at four male Sassos over one year).

However, the profit modeling in Table 15 does not capture the overall value of owning chickens. Household egg consumption is an important outcome that is not included in profitability calculations yet is relevant to the value of owning female chickens. To assess overall value, Table 16 repeats the analysis, replacing income from eggs sold with the total value of eggs produced.[47]

When using value of eggs produced rather than income from eggs sold, female ferenj chickens are more "valuable" than their habesha counterparts. Using averages, female ferenj chickens are more than 45 percent more valuable than female habesha chickens; using medians, they are over 100 percent more valuable. The difference between the value and profitability models for female chickens is likely driven by the fact that most household egg production goes towards household egg consumption rather than selling eggs at the market. Per this analysis, female ferenj chickens are more "valuable" than habesha chickens even if they are not more profitable at the one year mark.

Chicken Type	Statistic	Cost of Chick (ETB)	Value of Eggs Produced (ETB)	Income from selling chicken (ETB)	Feed Expenses (ETB)	All Other Expenses (ETB)	Value After Sale (ETB)
Female Habesha	Average	0	398.7	87.7	42.4	13.9	430.1
	Median	0	318.2	85	0	0	403.2
Female Ferenj	Average	76	704.0	160.7	118.8	45.1	624.8
	Median	76	751.4	153	0	0	828.4

Table 16: Per-chicken value model for female breeds (using data from direct effect and spillover samples)

Household Diet

Household nutrition is an additional key outcome of this study. We start by looking at total household food consumption, and then look individually at women and children's consumption.

Household food consumption is assessed by asking respondents which food items they and the members of their household have eaten in the last 7 days. Figure 19 shows that treatment households report consuming 3.0 more eggs in the previous week, compared to an average of 3.6 eggs in the control group. Treatment households also report consuming eggs 1.5 times per week, compared to .8 times per week in the control group. The results on egg consumption frequency and number of eggs consumed are slightly larger among the subset of households that reported no periods of non-normal food consumption (e.g.

[47] We calculate value of eggs produced per chicken by multiplying the total number of ferenj or habesha eggs produced by a household in one year with the most common price received for either type of egg and scaling this result by the number of female ferenj or habesha chickens. As before, the value of eggs produced is scaled by 0.85 to incorporate chicken mortality in the model.

fasting or holidays) in the previous week or in the previous month. The impact on household egg consumption—both in terms of the total number of eggs consumed per week and the weekly frequency of consumption—increases with ferenj flock size, as seen in the subgroup analysis of the appendix.[48]

The average household size for treatment and control groups is 6.2 and 6.3, respectively. Among control households, average egg consumption per household member per week is approximately 0.6 eggs. The average treatment effect is an additional 0.5 eggs per household member per week. To the extent that one egg is consumed by one member of the household, the increased egg consumption made possible by ferenj chickens will only benefit some members of the household and not others.

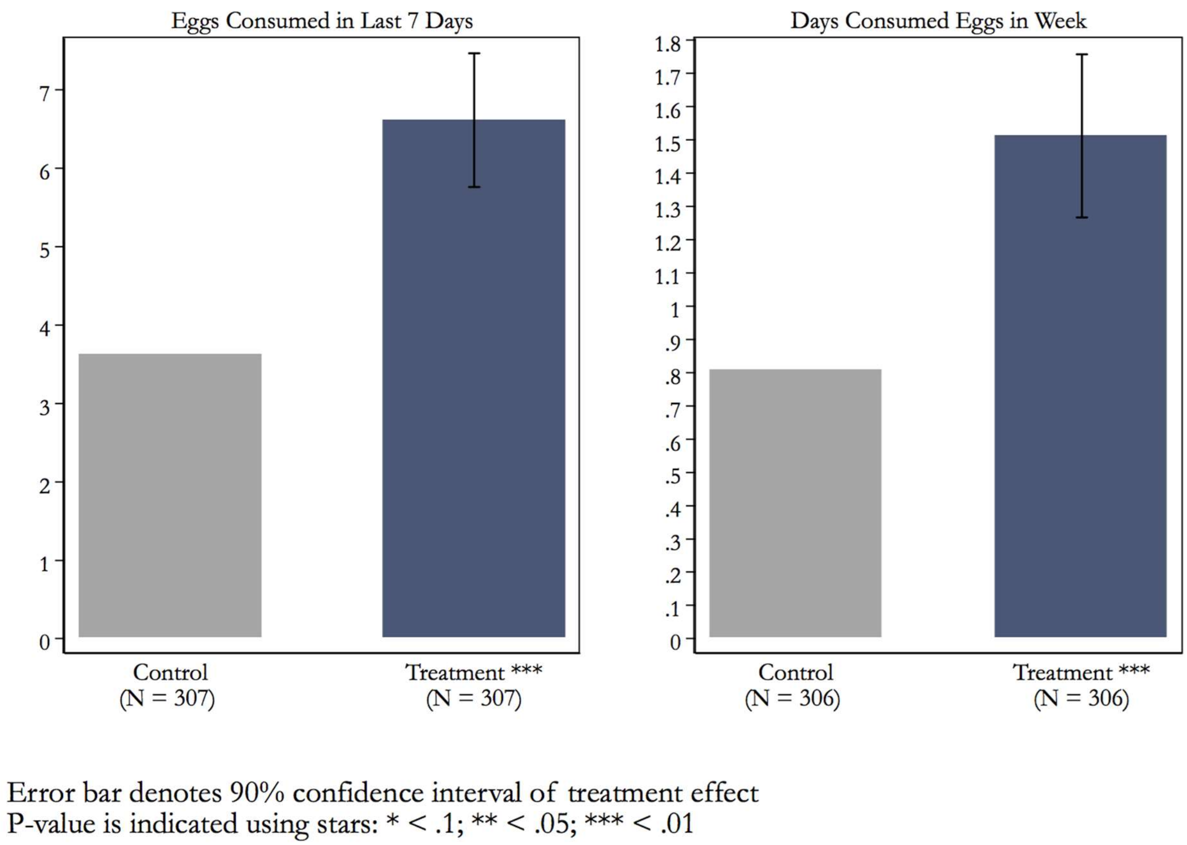

Figure 19: Impact estimate - household consumption of eggs

Figure 20 reports impact on the Food Consumption Score, which is a simple survey-based measurement of household food security developed by the World Food Programme.[49] We find that treatment households have increased their food consumption score, driven by increased consumption of eggs. Both treatment and control households are on average, above the threshold for an "acceptable" overall score

---

[48] This association between ferenj flock size and total household egg consumption does not seem to be due to increased egg consumption per chicken owned but rather constant egg consumption per chicken and an increase in the number of ferenj chickens owned.

[49] The FCS is a measure of food security "based on dietary diversity, food frequency, and relative nutritional importance." Per the technical guidance on the FCS, the maximum possible score is a 112, though values greater than 35 indicate an "acceptable" diet. World Food Programme VAM Team. *Food Consumption Analysis.* Rome: World Food Programme, 2008.

on the FCS. The results on egg consumption frequency and overall FCS score for households that reported no periods of non-normal consumption habits over the previous week or over the previous month are marginally larger in magnitude than the results reported for the overall sample. We investigated whether there were heterogeneous treatment effects among households that owned chickens at baseline and households that did not. The results from this analysis conform to the general pattern observed in the overall sample (see subgroup analysis in the appendix).

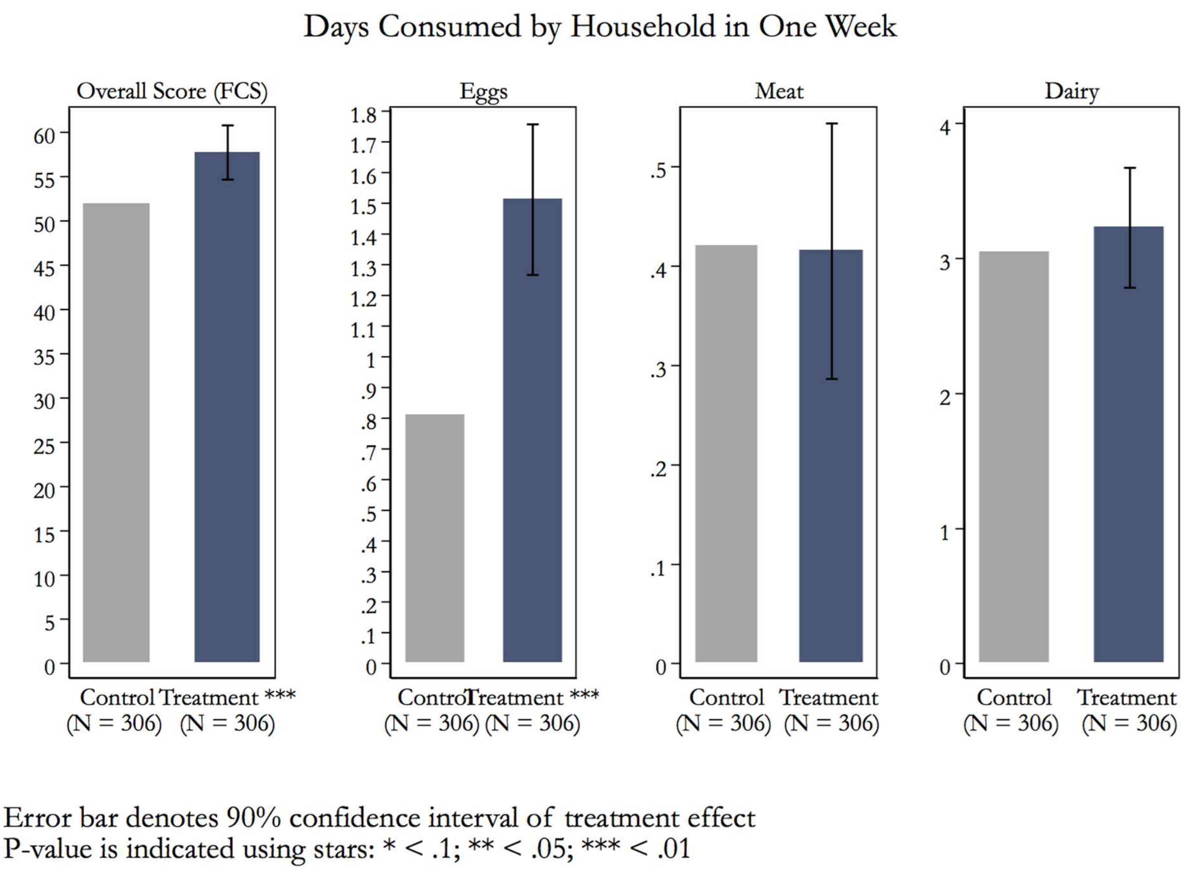

Figure 20: Impact estimate - food consumption score module

There is no significant increase in overall meat consumption; in fact, as seen in Figure 21 there is a slight decline in the number of chickens eaten in the last 6 months among treatment households. This may be due to the fact that treatment households have started viewing chickens more like a business, and therefore see chickens as a source of profit as opposed to consumption good. This business mentality may be reinforced by the fact that ferenj chickens command higher market prices. Households that did not own chickens at baseline appear to consume fewer chickens than households that did own chickens at baseline; this may be a consequence of chicken consumption occurring later in the chicken ownership cycle (i.e. after a chicken has reached consumable weight).

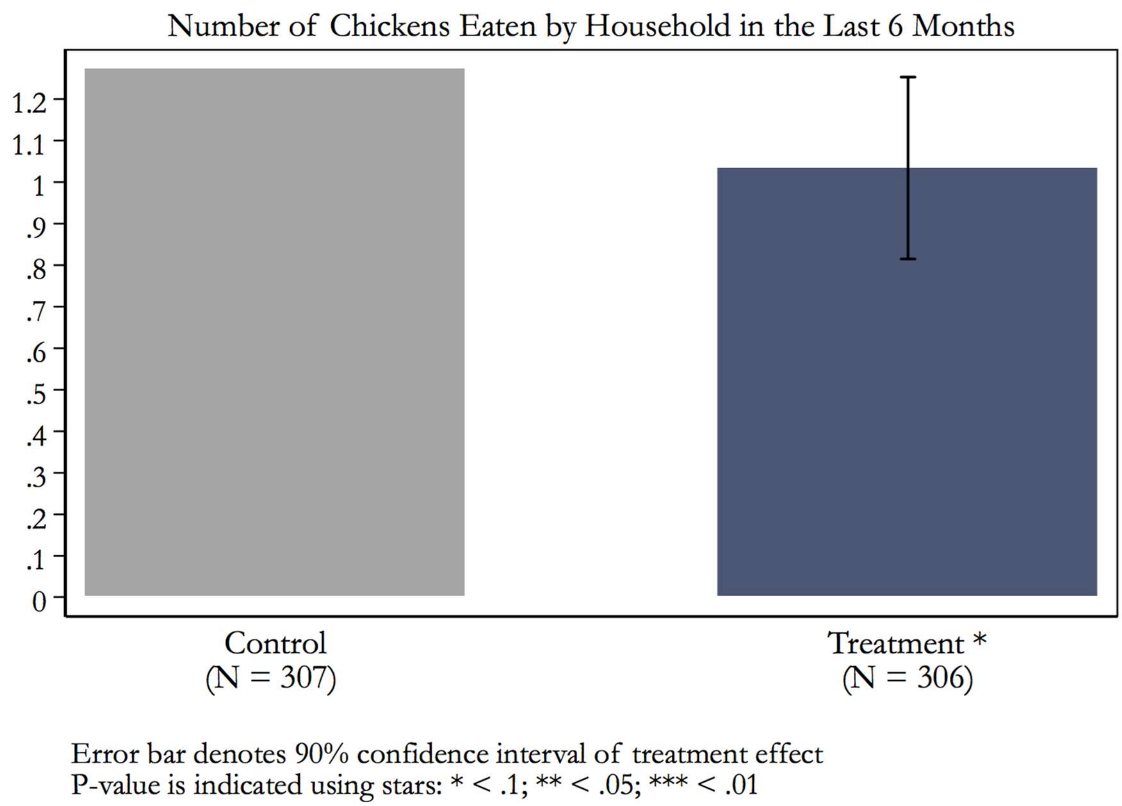

Figure 21: Impact estimate - chickens eaten by household

As shown in Figure 22, treatment households spend less than their control counterparts on egg expenditures. This makes sense because they now have more home-produced eggs.

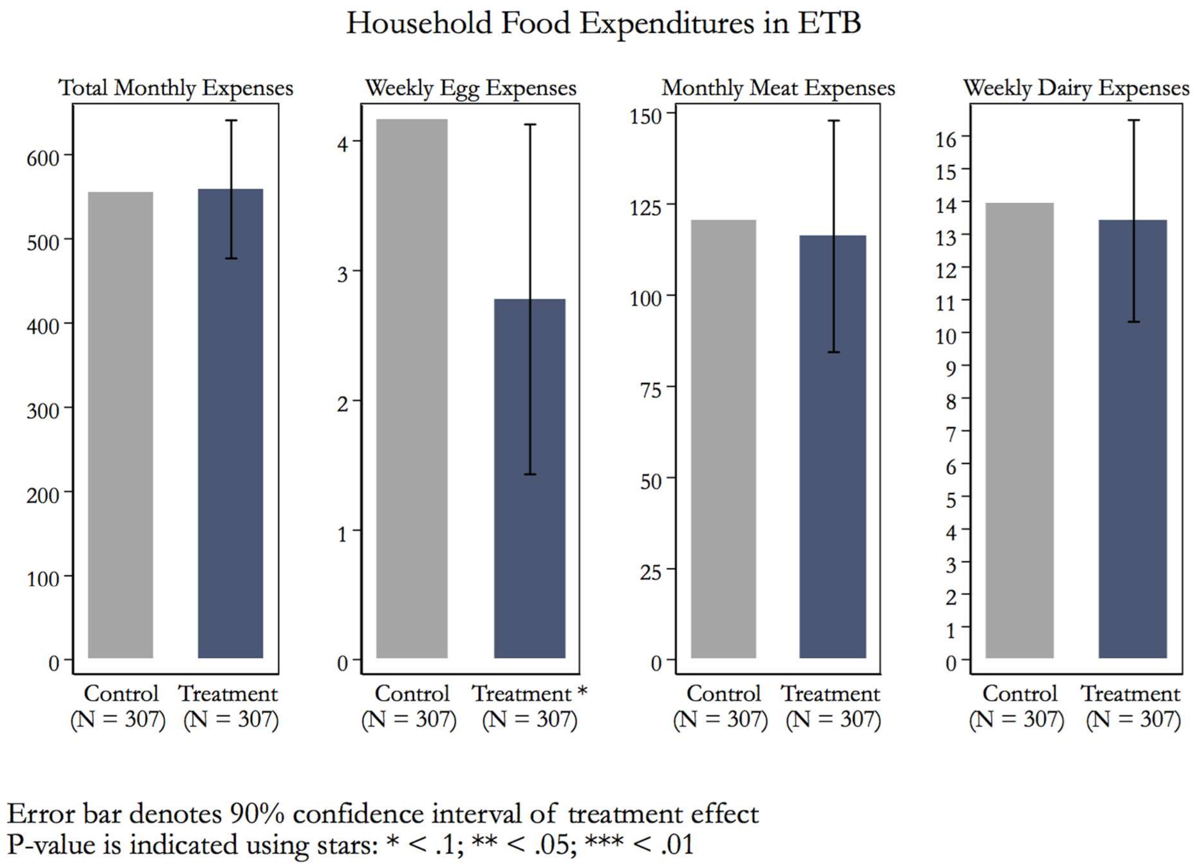

Figure 22: Impact estimate - household expenditures

Women's and Young Children's Nutrition

In addition to total household food consumption, we also asked detailed 24-hour food recall questions about the consumption of a woman and child under 5 in each household. We use this to look at consumption of specific food categories, as well as to construct the dietary diversity score (DDS). This is a simple proxy for nutritional quality developed by USAID.[50]

In Figure 23 we show the results of women's dietary diversity. We see a small, marginally significant increase in egg consumption, with 8.3% of women in treatment households reporting consuming eggs in the last 24 hours, compared to 4.9% of women in control households. The statistical significance of this result improves (to the 0.05 level) if we limit the analysis to women that reported no non-normal consumption habits over the previous 24 hours.[51] This effect on women's egg consumption is only

---

[50] The overall dietary diversity score is the total number of food groups consumed by the respondent in the 24 hour recall period. Swindale, Anne and Bilinsky, Paula. *Household Dietary Diversity Score (HHDS) for Measurement of Household Food Access: Indicator Guide.* Washington, D.C.: FANTA, 2006.
[51] If women did report non-normal food consumption in the previous 24 hours, they were asked to report their food consumption for "the most recent typical day." Only the statistical significance (indicating greater precision) and not the magnitude of the effect on women's egg consumption changes when the analysis is limited to women without non-normal food consumption in the previous day.

observed in households that owned ferenj chickens at or within 6 months of the endline data collection. Moreover, this effect appears to be driven by Muslim households and is less pronounced in Christian households.[52] Additionally, We do not see significant increases in consumption of meat, dairy, overall animal protein, or the dietary diversity score when looking at the overall sample.

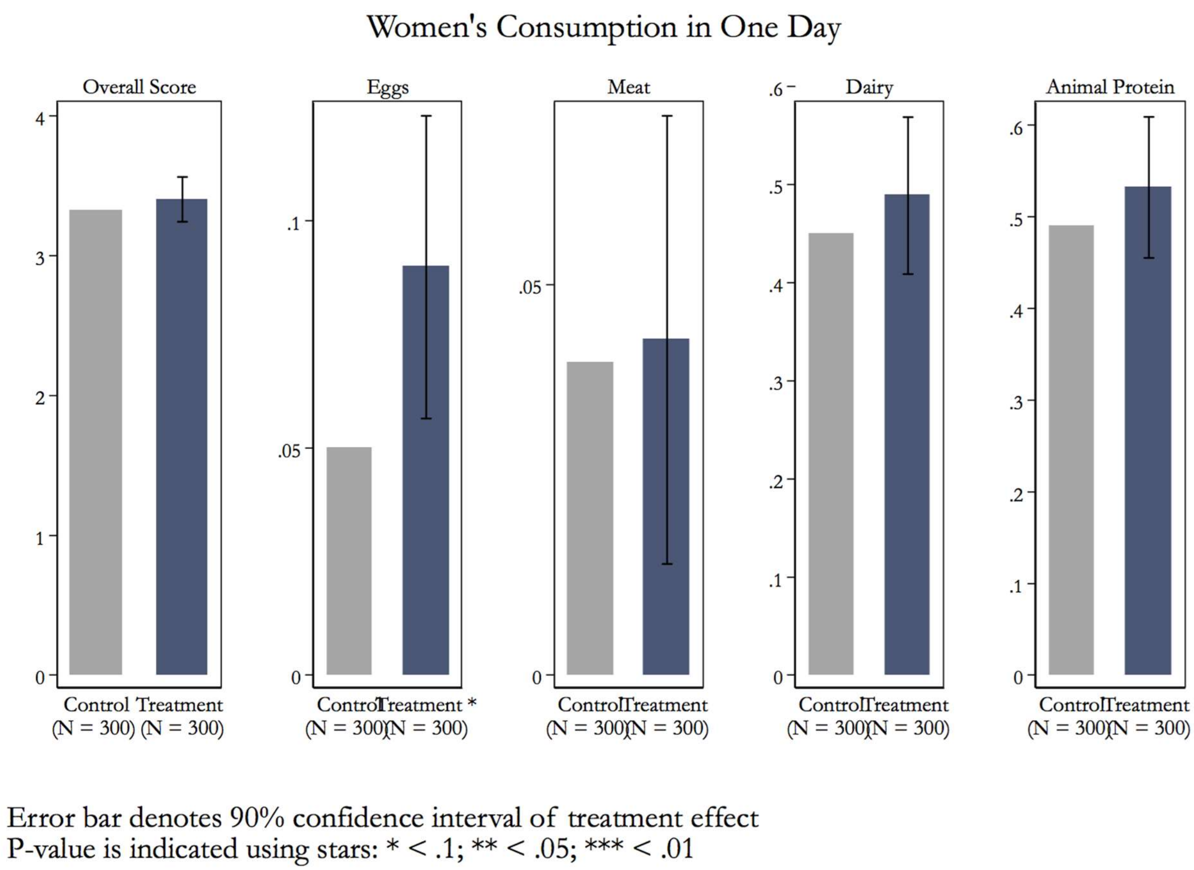

Figure 23: Impact estimate - women's dietary diversity

We also queried the dietary diversity of a child under 5 in the household. As shown in Figure 24, we do not see a significant increase of any consumption categories. We do see that children in the treatment group report higher consumption of dairy, though the precision of this result falls below standard cutoffs for statistical significance. It would be possible for chicken ownership to be increasing children's dairy consumption if households were using the increased income from eggs to purchase milk, or if households were bartering eggs for milk with their neighbors. This outcome seems unlikely, however, for the same reasons it would be unlikely for egg income to drive increased meat expenditures—the increase in income is just too minimal to entail meaningful increases in food expenditures. In association with

[52] However, Christian women were more likely to report that the previous day had been non-normal (e.g. a day of fasting) than Muslim women. If the previous day was non-normal, women respondents were asked about "the most recent typical day." To the extent women respondents did, in fact, report the most recent typical day, the heterogeneous treatment effect between Christian and Muslim households on women's egg consumption can be considered legitimate.

Acumen's Lean Data team, we conducted a follow-up phone survey with ferenj chicken owners to see if this link seemed plausible. Although some households reached as part of this effort reported using the money from selling eggs for food, no one mentioned purchasing dairy. (The most common purchases were oil, sugar, and salt.) Therefore we believe that the increase in dairy consumption is not an actual treatment effect of chicken ownership, and may be caused by imperfections in the matching process or statistical noise.

We also considered the same analysis separately for children under 2 and children from ages 2-5, but found similarly insignificant results for these subcategories. We did not explicitly ask about the food consumption of male household members; however, our results are consistent with increasing egg consumption for all members of the household with the exception of children.

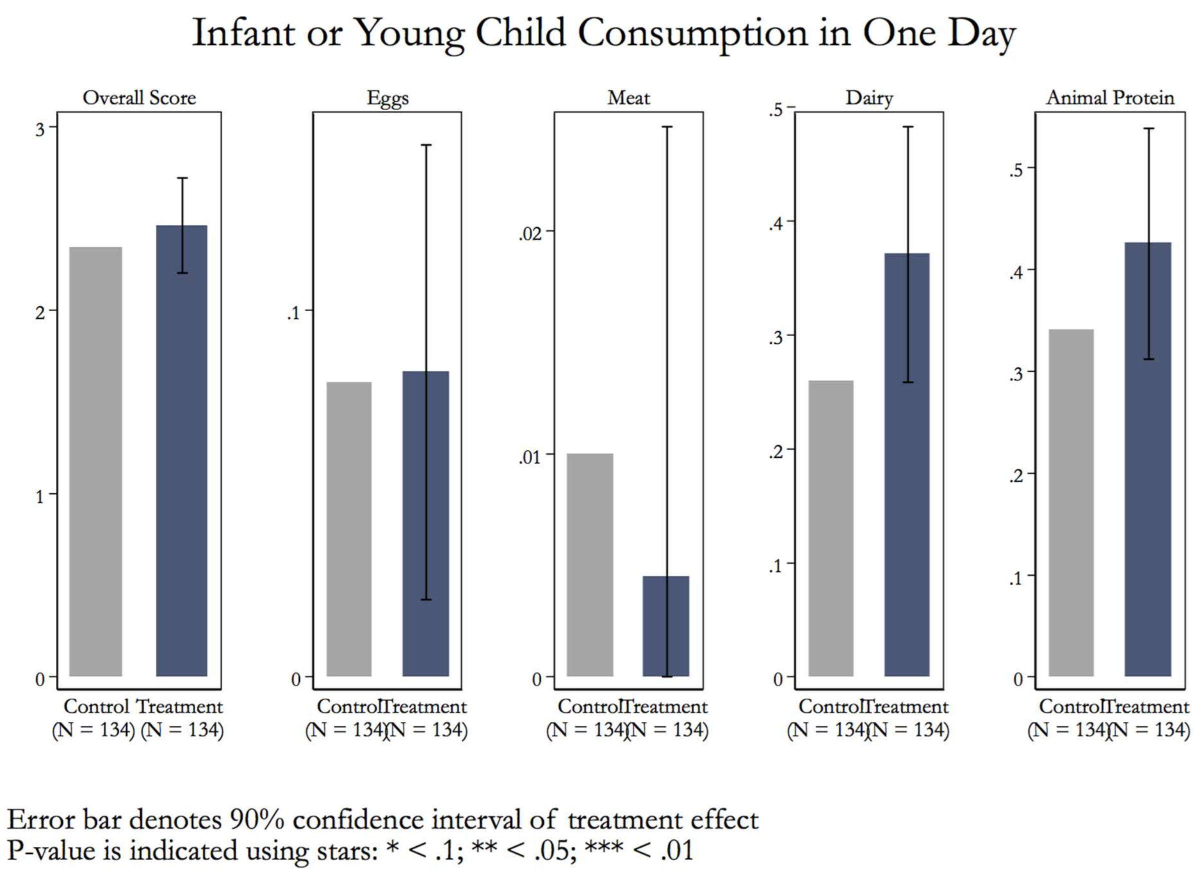

Figure 24: Impact estimate - child's food consumption

Women's Decision-Making

Since women tend to be in charge of poultry-rearing in Ethiopia, increasing income from poultry raising may raise womens' status in a household. We asked women a number of questions about whether they have decision-making power regarding poultry. This includes questions on whether they participated in buying chickens, participated in selling chickens and eggs, and had input into decisions on buying chickens, selling chickens, selling eggs, spending revenue from chicken and eggs, and the extent to which poultry products are included in the household diet. Questions regarding participation in buying chickens or selling chickens and eggs received "Yes" or "No" answers and were coded such that 1 corresponded to a "Yes" and 0 to a "No." For each of the other questions, the respond was asked to respond on a scale of 1-4, in which 1 was "I disagree very much" and 4 was "I agree very much" with the described activity.[53] To equalize the scale of all questions, responses from 1-4 were re-coded to 0-1 (taking on values of 0, 0.33 0.66, and 1). All the answers were then combined into an overall index, which is the average among all questions that applied to a specific household. Therefore a value of 1 on the index means that the woman has strong input into all decisions, and value of zero means she has limited input into any.

In Figure 25, we show that treatment households score higher on the index compared to control households, rising from .84 to 1.01.[54] Movement on the overall index is only seen in households that owned ferenj chickens at or within 6 months of endline data collection. Additionally, the treatment effect on the overall index is observed in Muslim households but not Christian households. However, the lack of movement on women's input on egg income and chicken income among the overall sample should give pause to any interpretation of higher productivity chickens being beneficial to women's empowerment. Among Muslim households and households that owned chickens at or within 6 months of endline, though, there is a significant increase in women's input on how egg income is used by 6 percentage points.

[53] As an example of an activity related to women's decision-making: "When your household sold eggs, did you get to decide how many eggs to sell, where to sell them, how much to sell them for, etc?"
[54] The smaller sample sizes found in Figure 25 are a result of the women's decision-making questions only being asked to households if the household had been involved in certain activities. For instance, women were only asked if they had input into the use of income from chicken sales if the household had actually sold a chicken. The same applies to women's input into the use of income from egg sales.

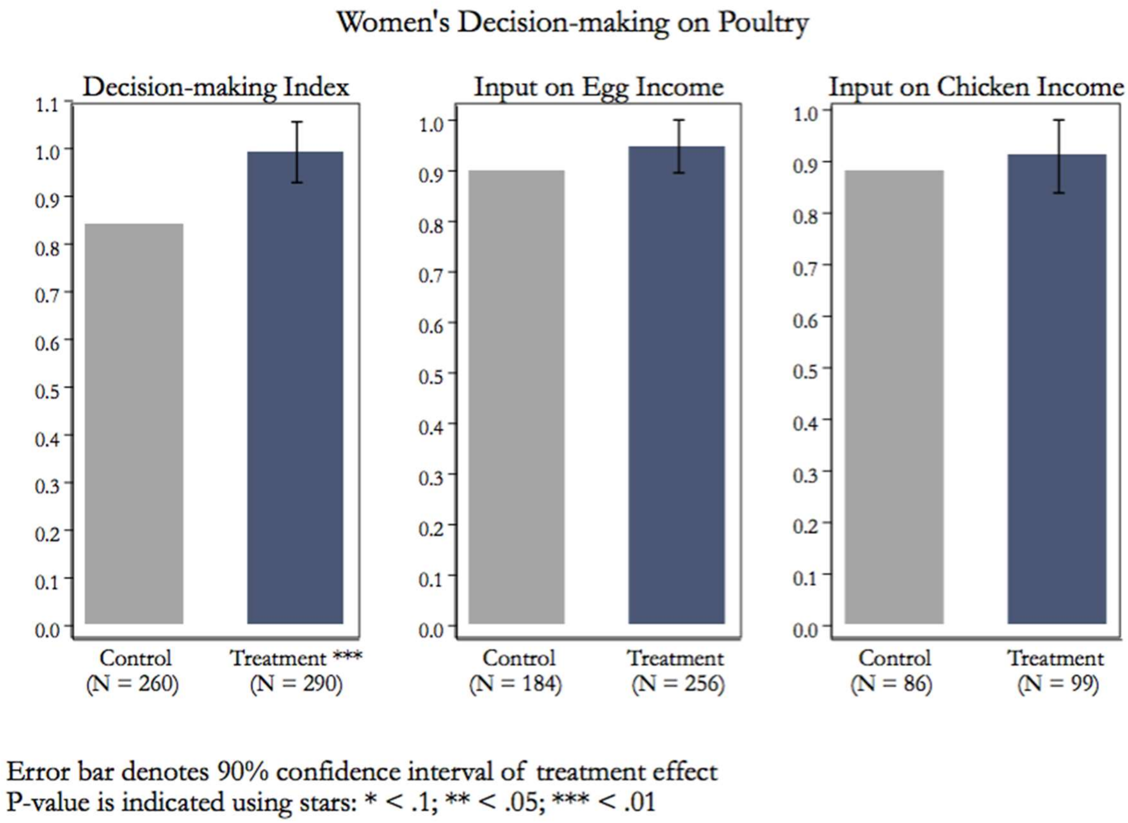

Figure 25: Impact estimate - women's decision-making on poultry

Spillovers

In theory, households that did not purchase chickens but lived near those who did may possibly also benefit from their neighbors' chicken ownership. However, given the marginal scale of the increase in egg production among treatment households, we would not expect to observe large spillover effects. In this section we assess potential spillovers by comparing non-purchasers who lived close to purchasers to other non-purchasers who did not live near ferenj purchasers.

The most obvious mechanism for spillovers would be that treatment households supply more eggs, which might increase egg consumption of their neighbors. As shown in Figures 26 and 27, we do not find any spillover effects on egg consumption, at both the household level and for women and children.

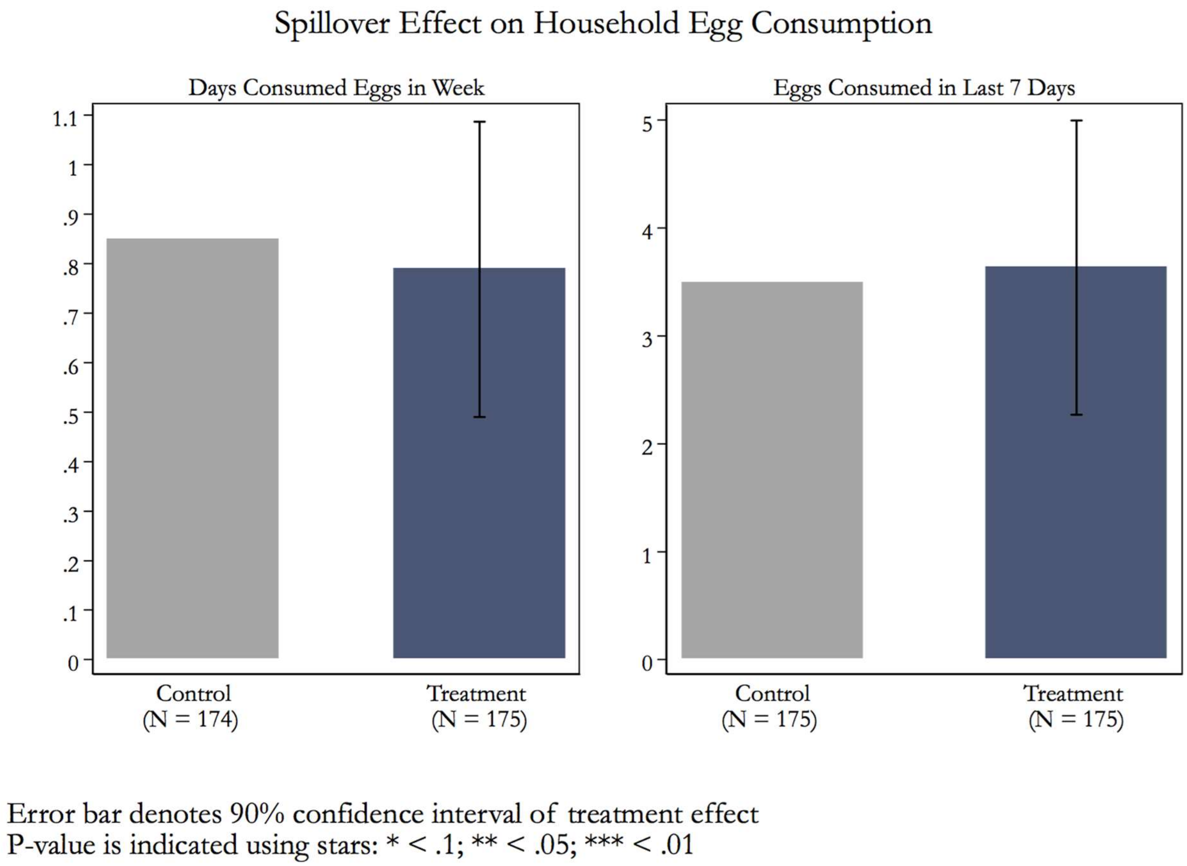

Figure 26: Spillover impact estimate - household egg consumption

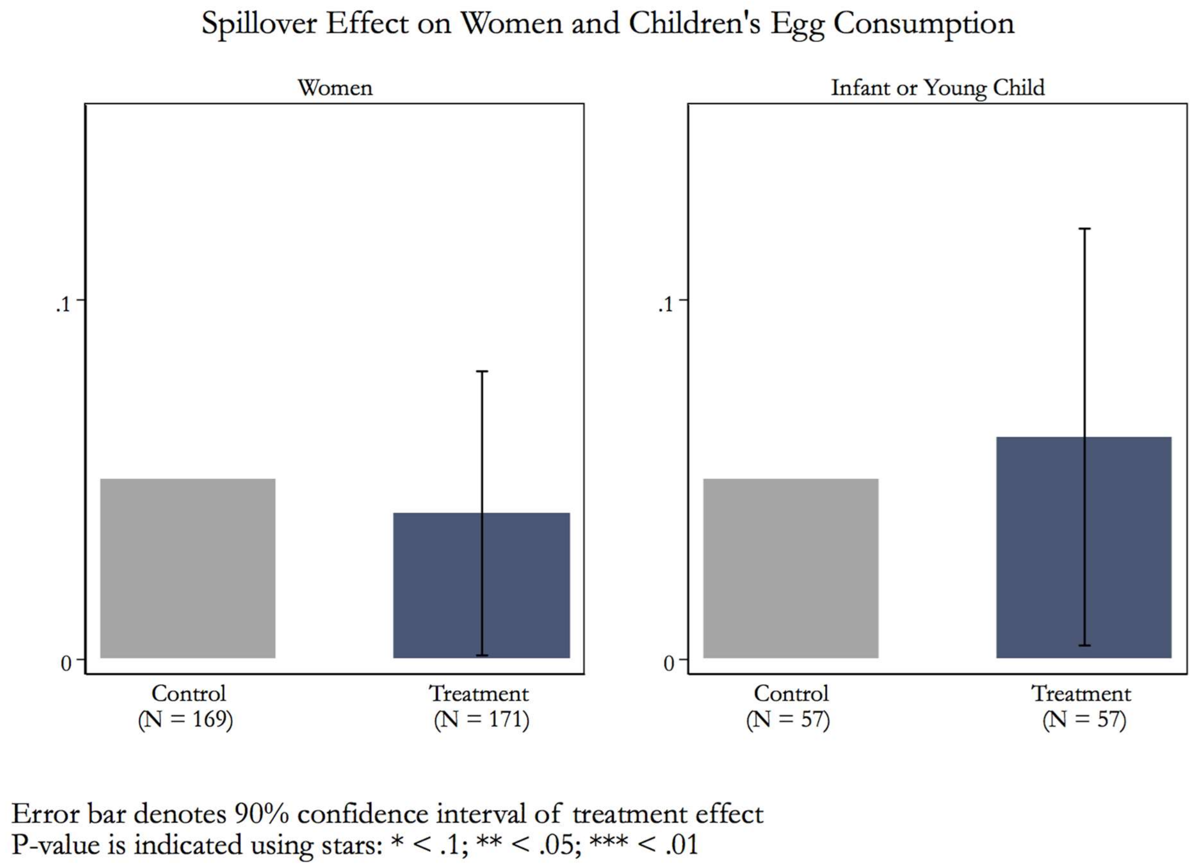

Figure 27: Spillover impact estimate - women's and children's egg consumption

We checked spillover results on other outcomes, and did not find any significant effects. Therefore we believe that spillovers do not represent an important path to impact in this setting where treatment households only owned, on average, two ferenj chickens.

Conclusions

This report finds evidence that backs up much of the theory of change, though with some important caveats. We find that ferenj chickens generate more eggs and command higher prices in the market, leading to both higher consumption of poultry products and increased revenue from their sale. Poultry revenue increases, but expenditure on ferenj chickens also increased, making long-term profitability of ferenj chickens uncertain. Net income over the last months was not higher for treatment households due to some household with very high expenses. However, net income did increase among households who had sold chickens in the last six months. Additionally, although ferenj chickens do not increase monthly profits on average, they do appear to be more “valuable” to households than their habesha counterparts when we use the market value of all eggs produced rather than income from eggs sold in the market.

We anticipate that the high level of expenditures on chicken-related activities in the treatment group are unlikely to have any positive linkage to chicken productivity. Households – particularly those in the treatment group – purchased some feed types (e.g. grains) that are not actually beneficial for chicken productivity. This may be a result of misinformation and households wishing to protect their investments in ferenj chickens with supposed higher quality inputs. This suggests that complimentary training or information campaigns could help ensure that raising ferenj chickens is profitable.

The outcomes on nutrition are also positive, though the practical significance of the results are muted. Increases in household nutrition are driven solely by egg consumption, though on average each member of the household is only able to consume an additional half egg per week. Importantly, we don't see a reported increase in animal consumption of children. It is possible that there are very small effects that this study is not powered to detect, but overall it is unlikely that ownership of improved chickens in their current form is leading to significant changes in nutrition of vulnerable household members. In order to achieve this objective, households might need larger flocks or additional behavioral change interventions.

It is important to note that this study comes after a relatively short period (6-14 months) of improved chicken ownership, and also among a population that on average had only a couple of improved chickens. It is possible that as household learn more about chicken ownership and increase their flock sizes that impacts will grow in the future. This is an area of future research, which can be at least partially explored through Acumen and EthioChicken's Lean Data collaboration.[55]

---

[55] Some households in our endline sample (for which there is information on flock size, expenses, and profitability) consented to being contacted in the future by EthioChicken. Following these households over time may provide insight into the evolution of flock size, expenses, and profitability.

Technical Appendix

Econometric Specifications

Baseline and endline data analysis was conducted using Stata 14.0 while elastic net modeling and matching were conducted using R. To estimate the direct effect of chicken purchases, we use the regression model specified in Equation 1, below, that regresses $treatment$ against the outcomes of interest. These outcomes reflect the core components of the causal chain found in the theory of change, including: egg production, egg sales, egg income, chicken sales, chicken income, households food consumption, household egg consumption, household meat consumption, women's dietary diversity, women's egg consumption, women's meat consumption, infant or young child (IYC) food consumption, IYC egg consumption, and IYC protein consumption.

$$outcome_i = \beta_0 + \beta_1\, treatment_i + \beta_2\, matching\ vars_i + \beta_3 holiday\ dummy_i \\ + \beta_4\, baseline\ measure_i + \varepsilon_i$$

Equation 1: Direct Effect Estimation Model

In this regression, i indexes each of the observations that are included in the regression. $\beta_1$is the estimated treatment effect, our main coefficient of interest. ε_i is an error term, denoting the difference between an observation's true value of an outcome measure and the regression model's prediction for that outcome. We assume that errors may be correlated within a kebele and are otherwise independent. We account for this by clustering standard errors at the kebele level in our main specification.

On the right-hand side of the equation, our regression includes a series of baseline variables used to identify the matched comparison households. Controlling for matching variables in a regression is standard practice for matching evaluations. These controls ($matching\ vars$) include: the probability of purchasing a ferenj chicken estimated using baseline data (see section on machine learning, below), as well as ferenj chicken ownership, flock size, chicken sales income, egg sale income, household egg consumption, and women's protein consumption. When regressing outcomes related to infant or young child feeding (IYCF), we also include the child's age and IYCF egg and protein consumption from baseline since they were used to match control and treatment households for those outcomes.

As mentioned in the Endline Sample section, surveying was halted due to unrest that occurred around the date of Ethiopian Orthodox Christmas. Due to our concern that data collected before and after the holiday might be different due to temporal sales and consumption patterns, we control for whether surveying took place before or after the holiday using the binary variable $holiday\ dummy$.

Lastly, whenever we have data on the baseline value for a particular outcome, we include baseline measures in our regression model. Including baseline values in the model improves the precision with which we can estimate the treatment effect of purchasing ferenj chickens. In this report, the majority of regression results include baseline controls but there are a minority of outcomes for which we do not have baseline information (specifically, we did not include the food consumption score in our baseline).

We use a similar model to estimate the spillover effect of proximity to ferenj chicken purchasers as seen in Equation 2. We use the same variables in matching spillover households with appropriate comparison households and therefore use the same variables as controls in our spillover regressions (Equation 2).

$$outcome_i = \beta_0 + \beta_1\, spillover\ treatment_i + \beta_2\, matching\ vars_i + \beta_3 holiday\ dummy_i + \beta_4\, baseline\ measure_i + \varepsilon_i$$

Equation 2: Spillover Effect Estimation Model

As noted in the Re-matching section of the report, the spillover "treatment" sample is composed of households with high exposure to other households that purchased ferenj chickens. This exposure metric is a sum of all nearby households that purchased ferenj chickens, weighted by distance. Matched controls are drawn from control kebeles.

Regression Tables

The main analyses discussed in the report are for the overall sample of 307 treatment and 307 control households, with covariate controls including the matching variables, baseline values for the outcome of interest, and whether data collection took place before or after the holidays. Standard errors for these results are clustered at the kebele level, since access to ferenj chickens was facilitated by kebele-level development agents. As robustness checks, we include additional model specifications in the appendix. Table 17 reports the treatment effect and its standard error for the main outcomes of interest with six different regression models:

1. Results with standard errors clustered at the kebele level – these are the results found in the report, and included here as a source of comparison with the other models
2. Results without clustered standard errors, but with Huber-White robust standard errors. It is not always standard in matching studies to cluster standard errors, so we include unclustered standard errors as a point of comparison.
3. Results with woreda fixed effects and standard errors clustered at the kebele level. We did not control for woreda effects in our main model: since treatment was strongly correlated with woreda, including woreda-level fixed effects decreases the contribution of certain woredas to the final estimate. However, we include these regressions as a robustness check.
4. Results with woreda fixed effects and unclustered (Huber-White robust) standard errors. We include woreda fixed effects for the reason outlined in (3) and unclustered standard errors for the reason discussed in (2).
5. Outcomes regressed on treatment only, with clustered standard errors. This allows us to look at the pure treatment effect rather than a treatment effect conditional on the inclusion of other regression covariates.
6. Outcomes regressed on treatment only and unclustered (Huber-White robust) standard errors. We uncluster standard errors for the reason discussed in (2).

Our results will be considered robust to model specification to the extent there is general agreement in the magnitude and significance of treatment effects across the different specifications. This is what we do, in fact, observe. For almost all of our major outcomes, the results from the different models are quite similar. The two exceptions are that our results for eggs sold and egg revenue lose some of their significance if outcomes are regressed only on the treatment. This indicates that there is a strongly identified treatment effect, conditional on the other variables included in the regression model.

	Results with Standard Errors Clustered at Kebele Level (results in report)	Results with Robust Standard Errors (not clustered)	Results with Woreda Fixed Effects and Standard Errors Clustered at Kebele Level	Results with Woreda Fixed Effects and Robust Standard Errors (not clustered)	Results without Control Variables and Standard Errors Clustered at Kebele Level	Results without Control Variables and Robust Standard Errors (not clustered)
Eggs Produced in 7 Days	6.60*** (0.87)	6.60*** (0.89)	7.67*** (1.18)	7.67*** (1.20)	6.78*** (0.98)	6.78*** (0.96)
Eggs Sold in 30 Days	4.49*** (1.66)	4.49*** (1.44)	3.86*** (1.40)	3.86*** (1.17)	1.98 (1.74)	1.98 (1.28)
Egg Revenue in 30 Days	14.38*** (5.32)	14.38*** (4.44)	12.04*** (4.43)	12.04*** (3.60)	6.69 (5.42)	6.69* (3.91)
Chickens Sold over 6 Months	0.41 (0.25)	0.41* (0.23)	0.67** (0.29)	0.67** (0.26)	0.50* (0.28)	0.50* (0.26)
Revenue from Chicken Sales over 6 months	88.05** (33.71)	88.05*** (31.40)	115.74*** (37.83)	115.74*** (34.59)	104.49*** (37.87)	104.49*** (36.18)
Monthly Chicken Income	16.84*** (3.78)	16.84*** (3.72)	17.43*** (4.15)	17.43*** (4.04)	12.89*** (4.55)	12.89*** (3.82)
Monthly Chicken Expenditures	28.48*** (4.15)	28.48*** (3.86)	32.44*** (4.96)	32.44*** (4.56)	28.11*** (4.36)	28.11*** (3.83)
Monthly Chicken Profit	-11.65** (5.69)	-11.65** (5.11)	-15.01** (6.03)	-15.01*** (5.58)	-15.22*** (5.65)	-15.22*** (4.84)
Days Consumed Eggs in Week	0.71*** (0.15)	0.71*** (0.13)	0.76*** (0.17)	0.76*** (0.15)	0.69*** (0.14)	0.69*** (0.13)
Eggs Consumed in Last 7 Days	2.99*** (0.52)	2.99*** (0.53)	2.97*** (0.59)	2.97*** (0.61)	2.94*** (0.57)	2.94*** (0.56)
Overall FCS Score	5.80*** (1.86)	5.80*** (1.63)	4.08** (1.86)	4.08** (1.77)	5.04** (2.07)	5.04*** (1.65)
Chickens Eaten by Household over 6 Months	-0.24* (0.13)	-0.24* (0.14)	-0.28* (0.16)	-0.28 (0.17)	-0.07 (0.18)	-0.07 (0.16)
Total Monthly Food Expenditures	3.21 (49.82)	3.21 (39.42)	-8.64 (50.12)	-8.64 (42.75)	-33.93 (58.11)	-33.93 (41.62)
Weekly Egg Expenses	-1.38* (0.82)	-1.38* (0.81)	-1.91** (0.90)	-1.91** (0.94)	-1.82** (0.90)	-1.82** (0.85)
Monthly Meat Expenses	-4.21 (19.29)	-4.21 (15.87)	-5.56 (20.23)	-5.56 (17.10)	-8.73 (21.30)	-8.73 (16.27)
Women's Dietary Diversity Score	0.08 (0.10)	0.08 (0.08)	0.02 (0.11)	0.02 (0.09)	0.06 (0.09)	0.06 (0.08)
Women's Egg Consumption	0.04* (0.02)	0.04* (0.02)	0.05** (0.02)	0.05** (0.02)	0.03 (0.02)	0.03* (0.02)
Women's Protein Consumption	0.04 (0.05)	0.04 (0.04)	0.01 (0.05)	0.01 (0.04)	0.00 (0.05)	0.00 (0.04)
Infant or Young Child Food	0.15 (0.16)	0.15 (0.16)	0.08 (0.19)	0.08 (0.18)	0.10 (0.16)	0.10 (0.16)

	Results with Standard Errors Clustered at Kebele Level (results in report)	Results with Robust Standard Errors (not clustered)	Results with Woreda Fixed Effects and Standard Errors Clustered at Kebele Level	Results with Woreda Fixed Effects and Robust Standard Errors (not clustered)	Results without Control Variables and Standard Errors Clustered at Kebele Level	Results without Control Variables and Robust Standard Errors (not clustered)
Consumption Score						
Infant or Young Child Egg Consumption	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.05)	0.01 (0.04)	0.01 (0.04)
Infant or Young Child Protein Consumption	0.09 (0.07)	0.09 (0.07)	0.09 (0.09)	0.09 (0.08)	0.11 (0.07)	0.11* (0.07)
Women's Decision-making Index	0.15*** (0.04)	0.15*** (0.04)	0.13*** (0.04)	0.13*** (0.04)	0.17*** (0.04)	0.17*** (0.04)
Women's Input on Egg Income	0.05 (0.03)	0.05 (0.03)	0.03 (0.03)	0.03 (0.03)	0.03 (0.02)	0.03* (0.02)
Women's Input on Chicken Income	0.03 (0.04)	0.03 (0.04)	0.03 (0.05)	0.03 (0.05)	0.02 (0.03)	0.02 (0.03)

Table 17: Robustness Checks - Various Model Specifications.
Standard errors (clustered at the kebele level) are included in parentheses.
*P-value is indicated using stars: * < .1; ** < 0.05; *** < 0.01*

In addition to these robustness checks, we also ran the main specification (control variables with clustered standard errors) for 10 different subgroups. In forming the subgroups, we only keep households that fit the required criteria for the subgroup as well as their matched comparisons from the control group (except for Muslim and Christian households). Note that the number of observations for the analysis of a particular treatment effect in the table may not be equal to the overall number of observations in a subgroup. For instance, the number of observations for the Infant and Young Child Feeding section has fewer observations for all subgroups because not every household in each subgroup has a child under the age of five. We include the following subgroups in Table 18:

- Column 1: the overall sample as a point of reference
- Column 2: households that owned more than 5 ferenj chickens at the time of endline data collection. Our intention with this subgroup analysis, as with the following subgroup, is to investigate whether there are heterogeneous treatment effects by ferenj flock size. It is important to note, though, that ferenj flock size at endline is correlated with treatment: group sorting is in part a reflection of treatment status that may limit the interpretability of heterogeneous effects.
- Column 3: households that owned more than 4 ferenj chickens at the time of endline data collection.
- Column 4: Christian households. Our overall sample contained a higher percentage of Muslim households than the general population in Ethiopia so we wanted to assess variation in treatment effects between households of the two different religious traditions. Creation of this subgroup

was based on baseline data. A household's religion was not incorporated into the matching algorithm; while we can look at the subset of households in treatment and control categories that identified as Christian or Muslim, we cannot look at the subset of matched pairs that were either Muslim or Christian due to a dramatic resulting loss in sample size. Looking at subgroups by religion may "undo" some of the matching algorithm in the sense that the treatment and control groups are no longer even and may vary in terms of their other underlying characteristics.

- Column 5: Muslim households. Our motivation for analyzing this subgroup is the same as above.
- Column 6: Households that owned ferenj chickens at endline. As found in this report, selling chickens is an important component of overall income streams from chickens. We wanted to include households that had already sold their ferenj chickens in our overall sample so as to capture benefits from chicken sales. However, the impact of many outcomes may be stronger if the sample is limited to households that actually own chickens at the time of endline data collection.
- Column 7: Households that did not own ferenj chickens at endline. As above, we wanted to analyze whether there were heterogeneous treatment effects on the basis of ferenj chicken ownership at endline.
- Column 8: Households that owned ferenj chickens at or within 6 months of endline data collection. Similar to the rationale of the subgroup for ferenj chicken owners at endline, this subgroup allows us to look at impacts on households that either currently own ferenj chickens or had gotten rid of their chickens recently, thereby allowing us to capture some degree of revenue from chicken sales.
- Column 9: Households that did not own ferenj chickens at or within 6 months of endline data collection. This subgroup analysis allows us to gain some leverage over whether impact is confined to households that currently own ferenj or that recently sold them.
- Column 10: Households that owned chickens at baseline. This subgroup was motivated by the desire to look at whether patterns of food consumption differed for households that owned chickens prior to the intervention compared to households that did not. We did not exactly match on baseline flock size, so 30 observations that owned chickens at baseline (14 treatment and 16 control) were dropped from this subgroup analysis due to the fact they were matched with a household that did not own chickens at baseline.
- Column 11: Households that did not own chickens at baseline. This subgroup analysis allows us to see the other side of the results in Column 10. As above, 30 observations that did not own chickens at baseline (16 treatment and 14 control) were dropped from this subgroup analysis due to the fact they were matched with a household that did own chickens at baseline.

To the extent that a result is consistent across samples, we can make the claim that a result appears generalizable for the entire sample rather than a result specific to a certain subgroup. As seen in the table, there is broad agreement on the magnitude and statistical significance of treatment effects, suggesting that our main results (egg production, profits from chicken income, egg consumption) are robust to subgroup analyses.

	Total Sample	HHs with more than 5 FJ chickens	HHs with more than 4 FJ chickens	Christian HHs	Muslim HHs	HHs that owned FJ at endline	HHs that did not own FJ at endline	HHs that owned FJ at or within 6 months of endline	HHs that did not own FJ at or within 6 months of endline	HHs that owned chickens at baseline	HHs that did not own any chickens at baseline
Number of observations in subgroup analysis	614 (307 T & 307 C)	46 (23 T & 23 C)	76 (38 T & 38 C)	134 (75 T & 59 C)	478 (232 T & 246 C)	392 (196 T & 196 C)	222 (111 T & 111 C)	438 (219 T & 219 C)	176 (88 T & 88 C)	404 (202 T & 202 C)	150 (75 T & 75 C)
Outcome	*Treatment Effect, by Subgroup*										
Eggs Produced in 7 Days	6.60*** (0.87)	24.59*** (6.19)	19.6*** (4.27)	5.16** (1.99)	6.81*** (1.02)	9.40*** (1.17)	1.78* (1.03)	8.03*** (1.06)	3.44*** (1.12)	6.81*** (1.01)	7.35*** (1.93)
Eggs Sold in 30 Days	4.49*** (1.66)	7.01 (5.20)	5.00 (3.75)	4.01* (2.12)	4.59** (1.86)	5.70*** (2.03)	1.58 (1.30)	5.39*** (1.92)	1.68 (1.32)	5.20** (2.22)	3.63** (1.49)
Egg Revenue in 30 Days	14.38*** (5.32)	23.13 (16.87)	14.62 (11.17)	14.49* (7.49)	14.31** (5.71)	18.11*** (6.36)	5.59 (4.69)	17.13*** (6.04)	5.89 (4.74)	16.96** (7.13)	11.08** (4.81)
Chickens Sold over 6 Months	0.41 (0.25)	1.71 (1.02)	1.35* (0.74)	1.50** (0.64)	0.08 (0.25)	0.42 (0.28)	0.44 (0.45)	0.48* (0.29)	0.34 (0.56)	0.27 (0.26)	0.46 (0.40)
Revenue from Chicken Sales over 6 months	88.05** (33.71)	319.72* (161.34)	271.85** (110.82)	280.66*** (92.15)	29.14 (31.18)	93.23** (38.71)	83.01 (57.87)	107.76*** (39.01)	42.76 (66.67)	78.76** (37.96)	91.98 (56.70)
Monthly Chicken Income	16.84*** (3.78)	42.64*** (12.50)	37.35*** (10.91)	31.75*** (8.27)	11.75*** (3.85)	20.26*** (4.26)	9.77 (6.37)	21.37*** (4.22)	4.46 (5.77)	16.84*** (4.38)	21.37*** (8.04)
Monthly Chicken Expenditures	28.48*** (4.15)	99.77*** (19.89)	91.6*** (17.02)	22.91*** (8.18)	30.29*** (4.80)	38.18*** (5.69)	9.88** (3.99)	35.53*** (5.10)	9.44* (4.76)	30.21*** (5.37)	24.53*** (8.13)
Monthly Chicken Profit	-11.65** (5.69)	-60.22** (22.48)	-53.16*** (18.08)	8.91 (10.81)	-18.57*** (5.92)	-17.74** (7.09)	0.04 (7.13)	-14.04** (6.65)	-4.74 (6.80)	-13.31** (6.65)	-3.25 (12.04)

	Total Sample	HHs with more than 5 FJ chickens	HHs with more than 4 FJ chickens	Christian HHs	Muslim HHs	HHs that owned FJ at endline	HHs that did not own FJ at endline	HHs that owned FJ at or within 6 months of endline	HHs that did not own FJ at or within 6 months of endline	HHs that owned chickens at baseline	HHs that did not own any chickens at baseline
Days Consumed Eggs in Week	0.71*** (0.15)	1.62*** (0.50)	1.31*** (0.38)	0.14 (0.31)	0.84*** (0.16)	0.98*** (0.18)	0.24 (0.17)	0.90*** (0.17)	0.25 (0.20)	0.66*** (0.17)	0.99*** (0.26)
Eggs Consumed in Last 7 Days	2.99*** (0.52)	6.95*** (1.84)	6.27*** (1.33)	2.83** (1.32)	2.87*** (0.60)	3.91*** (0.69)	1.34* (0.79)	3.42*** (0.68)	1.92** (0.81)	3.24*** (0.73)	2.94*** (1.06)
Overall FCS Score	5.80*** (1.86)	7.54 (7.67)	5.92 (4.19)	6.88 (4.25)	5.41*** (2.05)	7.39*** (2.19)	2.57 (2.90)	7.10*** (2.11)	1.86 (3.19)	6.53*** (2.07)	4.51 (3.79)
Chickens Eaten by Household over 6 Months	-0.24* (0.13)	0.01 (0.55)	0.19 (0.43)	-0.56 (0.49)	-0.15 (0.14)	-0.28 (0.18)	-0.13 (0.25)	-0.24 (0.16)	-0.20 (0.31)	-0.06 (0.16)	-0.50* (0.25)
Total Monthly Food Expenditures	3.21 (49.82)	129.44 (156.10)	97.25 (150.60)	14.27 (81.46)	-10.09 (58.74)	-8.66 (52.23)	22.52 (75.07)	13.23 (54.42)	-38.10 (69.74)	16.94 (51.77)	-25.55 (85.36)
Weekly Egg Expenses	-1.38* (0.82)	-6.60** (2.66)	-4.20** (1.84)	-2.12 (2.02)	-0.95 (0.90)	-2.43*** (0.87)	0.51 (1.78)	-1.38 (0.94)	-1.64 (1.45)	-0.88 (1.00)	-3.19* (1.84)
Monthly Meat Expenses	-4.21 (19.29)	21.45 (87.69)	32.24 (65.24)	6.24 (46.21)	-14.16 (20.70)	-6.22 (24.75)	-1.51 (25.88)	-1.16 (22.89)	-13.38 (25.94)	2.46 (2.46)	1.09 34.60
Women's Dietary Diversity Score	0.08 (0.10)	-0.48 (0.34)	-0.40 (0.26)	0.14 (0.22)	0.05 (0.11)	0.00 (0.11)	0.19 (0.17)	0.03 (0.11)	0.19 (0.20)	0.11 (0.11)	0.11 (0.19)
Women's Egg Consumption	0.04* (0.02)	0.12 (0.10)	0.07 (0.06)	0.01 (0.06)	0.03* (0.02)	0.05* (0.03)	0.01 (0.03)	0.05** (0.02)	0.01 (0.03)	0.05* (0.03)	0.04 (0.03)
Women's Protein Consumption	0.04 (0.05)	-0.01 (0.17)	-0.06 (0.14)	0.02 (0.10)	0.05 (0.05)	0.07 (0.05)	-0.01 (0.08)	0.07 (0.05)	-0.06 (0.09)	0.07 (0.05)	0.01 (0.09)

	Total Sample	HHs with more than 5 FJ chickens	HHs with more than 4 FJ chickens	Christian HHs	Muslim HHs	HHs that owned FJ at endline	HHs that did not own FJ at endline	HHs that owned FJ at or within 6 months of endline	HHs that did not own FJ at or within 6 months of endline	HHs that owned chickens at baseline	HHs that did not own any chickens at baseline
Infant or Young Child Food Consump-tion Score	0.15 (0.16)	-0.30 (0.31)	0.25 (0.70)	0.38 (0.45)	0.08 (0.17)	0.10 (0.20)	0.18 (0.29)	0.18 (0.20)	0.11 (0.33)	0.32* (0.19)	0.03 (0.31)
Infant or Young Child Egg Consump-tion	0.00 (0.04)	0.31 (0.39)	0.37 (0.26)	-0.43 (0.25)	0.01 (0.04)	0.02 (0.06)	0.00 (0.06)	0.03 (0.05)	-0.05 (0.07)	0.05 (0.05)	-0.08 (0.08)
Infant or Young Child Protein Consump-tion	0.09 (0.07)	0.32 (0.72)	0.33 (0.23)	0.05 (0.20)	0.07 (0.08)	0.17* (0.09)	-0.03 (0.11)	0.16* (0.08)	-0.09 (0.13)	0.15* (0.09)	0.05 (0.05)
Women's Decision-making Index	0.15*** (0.04)	0.10 (0.11)	0.16 (0.10)	0.00 (0.07)	0.19*** (0.05)	0.19*** (0.04)	0.07 (0.07)	0.20*** (0.04)	0.03 (0.07)	0.14*** (0.04)	0.25** (0.12)
Women's Input on Egg Income	0.05 (0.03)	0.01 (0.02)	0.12 (0.11)	-0.02 (0.03)	0.06* (0.04)	0.08** (0.03)	-0.01 (0.05)	0.08** (0.04)	-0.08 (0.05)	0.02 (0.03)	Too Few Obs.
Women's Input on Chicken Income	0.03 (0.04)	Too Few Obs.	Too Few Obs.	0.09 (0.06)	0.02 (0.05)	0.08* (0.04)	-0.03 (0.05)	0.06 (0.04)	-0.07 (0.05)	0.01 (0.05)	Too Few Obs.

Table 18: Subgroup Analyses

Standard errors (clustered at the kebele level) are included in parentheses.

*P-value is indicated using stars: * < .1; ** < 0.05; *** < 0.01*

Machine Learning: Predicting the Probability of Purchasing Ferenj Chickens

Data from baseline provided a wealth of information on the pre-treatment characteristics of households that ended up purchasing ferenj chickens. We used machine learning to leverage this information for the purposes of matching treated households with appropriate comparison households by calibrating a model of chicken purchasing that predicts which households ultimately purchase. This machine learning model was trained on households in treatment kebeles before being used on households on control kebeles (kebeles without access to ferenj chickens) to determine which households would likely have purchased ferenj chickens had they been able.

The machine-learning model used for this prediction was an elastic net. These models are among the most frequently used in data science applications and are related to normal linear regressions. The problem with normal linear regressions in this context is that—with a multitude of possible explanatory variables and minimal prior information about which variables are most important for determining ferenj chicken purchasing—they tend to overfit the model (modeling a sample's data so well that the model can't be generalized outside of the sample). The elastic net operators fix this problem by decreasing the values of estimated regression coefficients (dubbed "shrinkage" in the machine learning literature) and leaving regressors out of the model if they lack sufficient explanatory power ("selection").

There are two parameters that define an elastic net procedure: alpha and lambda. Alpha determines the balance between shrinkage and selection while lambda determines the magnitude of the shrinkage or selection. A "ridge" regression maximizes shrinkage while a "lasso" regression maximizes selection. For our elastic net model, we chose alpha equal to one, which signifies a lasso regression. We decided to maximize selection because we believed that generating a more parsimonious model of purchasing propensity would be the most useful for EthioChicken.

Lambda is determined through a process known as cross-validation. This means that the model is trained on many different random draws of the dataset and then the model's performance is assessed on the remainder of the dataset. This process identifies the lambda that works best in the algorithm. We trained the elastic net using R's glmnet package with default options, including cross-validation with 10 folds. This algorithm chose a value of lamba equal to .0214, and selected 39 variables from the baseline data to predict chicken purchasing. Table 19 below lists the variables chosen by the lasso regression and their weights. The weights can be interpreted in the same manner as regression coefficients: the contribution of a marginal change in a particular input to the predicted value.

Variable Name	Weight
Intercept	0.232
Head of household completed some secondary school	0.012
Head of household completed a diploma or certificate program	0.061
If Household is Orthodox Christian	0.017
Progress out of Poverty Index: Jewelry Ownership	0.018
Unit of Land Ownership	0.039
If Primary Crop is Barley	0.073

Variable Name	Weight
If Primary Crop is Khat	-0.014
Ox Ownership	0.021
Ferenj Cow Ownership	0.077
Crossbreed Cow Ownership	0.023
Sheep Ownership	0.004
If Household Did Not Own Chickens but Tried to Purchase Chickens in the Last 12 Months	0.048
Number of Habesha Chickens that Were Stolen	0.014
If Household Had Productive Habesha Chickens all months in the previous year	0.029
If Household Had Productive Habesha Chickens for 8 Months out of the last year	0.012
The Number of Habesha Eggs Sold in a Typical Week	0.001
Whether Household Owned Ferenj Chickens for at Least One Year	0.088
Number of Ferenj Male Dual Purpose Chickens Currently Owned	0.0004
Number of Ferenj Female Dual Purpose Chickens Owned in Last 6 Months	0.033
Whether Household Consumed Ferenj Chickens in the Last 6 Months	0.053
If Household Specified “Other” to Source of Remaining Eggs Consumed	0.057
Average Household Egg Consumption Per Week over Prior 6 Months	0.002
Whether Household Spent Money on Vaccines for Chickens	0.114
If Respondent Thinks Investing in Ferenj Chickens is Safe	-0.00003
Whether Previous 30 Days were Normal Periods of Food Consumption	0.0005
Infant and Young Child Consumption of Nuts	-0.004
Women’s Consumption of Organ	-0.117

Variable Name	Weight
Meat	
Women's Consumption of Poultry	0.085
Women's Consumption of Vegatables	-0.0005
Poverty Score	0.0004
If Agriculture is a Source of Household Income	-0.015
If Household Owns at Least One Ferenj Dual Purpose Male Chicken	0.008
Price Received for One Ferenj Chicken	0.00009
Whether Household Sold their Habesha Chickens for Agricultural Products	0.018
Types of Chickens Owned in the Last 6 Months	0.006
Total Eggs Given Away	0.002
If Household Stopped Raising Chickens Because of Chicken Misbehavior	-0.037
If Household Stopped Raising Chickens Because of Other Reasons	0.00006
Women's Consumption of Meat and Fish	0.039

Table 19: Variables and Weights Chosen by Lasso Regression

We need to take caution in not over-interpreting the variables and weights identified in our elastic net model: we cannot say whether a particular variable *caused* households to purchase ferenj chickens. For instance, we may observe that the number of male dual-purpose ferenj chickens that a household currently owns is positively correlated with additional purchases of ferenj chickens. However, this is not sufficient to claim that current male dual-purpose ferenj chicken ownership will lead to more households purchasing additional ferenj chickens. It may be that current male dual-purpose ferenj ownership is correlated with one or several other variables that are driving the purchase of ferenj chickens, such as household income, access to ferenj chickens, or other factors. The prediction algorithm cannot distinguish between which variables are true 'drivers' versus 'proxies', and if two highly correlated variables exist in the data, then the algorithm will essentially randomly choose which one to include. For this reason, any interpretation of the specific variables used in a prediction algorithm is only suggestive.

Figure 28, below, shows a histogram of the predicted values for purchasing a ferenj chicken among households that did in fact purchase (blue bars) and among households that did not in fact purchase (red bars). As we would expect, households that ended up actually purchasing ferenj chickens have higher predicted probabilities of purchasing, on average, than households that did not purchase.

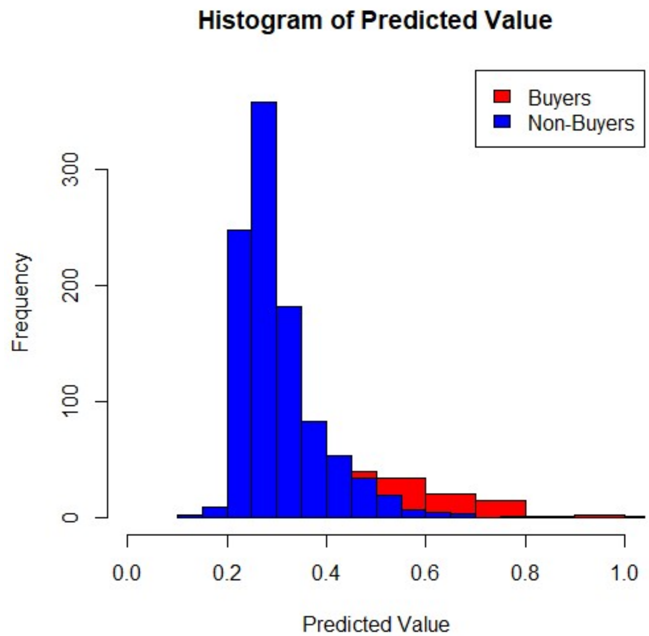

Figure 28: Histogram of Predicted Values

Unfortunately, doing 1:1 matching based on the lasso model's predicted probability of purchase gave poor balance on other key variables that we believed were important. This is likely due to the fact that treatment and control kebeles were not geographically balanced. Therefore, we used the predicted value as one variable in a more comprehensive matching algorithm.

Matching Procedure

We use genetic matching to find appropriate matches between treatment and control households, using baseline characteristics. Genetic matching is a method that uses an iterative algorithm to attempt to find matches that provide good balance on all variables simultaneously[56]. To conduct the genetic matching, we used the MatchIt package in R. This program allows the user to determine the closeness of matches along each of the matching covariates by setting calipers. The calipers indicate how close treatment and control matches need to be (in standard deviations) for the match to be considered high-quality. A caliper of zero will require exact matches on a variable, small calipers will require close matches, and large calipers will allow for greater distance. The calipers we used for each variable in the genetic matching algorithm are included in Table 20 below.

[56] Diamond, Alexis, and Jasjeet S. Sekhon. "Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies." *Review of Economics and Statistics* 95.3 (2013): 932-945.

Matching Covariate	Caliper for Direct Effect Matching (in standard deviations)
Distance (similar to propensity score)	1
Predicted Value of Purchasing	1
Current FJ Owner	0
Owned FJ One Year	2
FJ Current Flock Size	2
Eggs Eaten in Last Week	2
Egg Income in Last Month	2
Chicken Income in last 6 months	1.5
Total Current Flock	2
Woman consumed protein	2

Table 20: Calipers used for Direct Effect Matching

As discussed in the Re-matching section of the report, we ran an additional matching algorithm specifically for infant and young child feeding (IYCF) outcomes. This allowed us to ensure balance on child-specific variables by incorporating baseline values of child nutrition as matching covariates in addition to the previous matching variables. However, some households that reported children under 5 at endline did not have children under 5 at baseline. Omitting children present at endline but not baseline may have biased our results, yet we could not match on baseline child-specific values for these households. Instead, we conducted two matches: one for households that reported children at baseline and for which we could incorporate baseline child nutrition information, and a second for households that only reported children at endline and that were not matched using baseline child nutrition data. Table 21 shows the variables that were included in these matches and the calipers used. All matching variables are from baseline unless otherwise specified.

Matching Covariate	Caliper for IYCF Matching, Households with Children at Baseline	Caliper for IYCF Matching, Households without Children at Baseline
Distance	1.5	1.5
Predicted Value	1.5	1
Child Under 2 Available (at endline)	0	0
Child Under 5 Available (at endline)	0	0
Child's Age (at endline)	1.75	2
IYCF Egg Consumption	2	Not included
IYCF Chicken Consumption	2	Not included
IYCF Protein Consumption	2	Not included
Current FJ Owner	0	0
Owned FJ One Year	2	2
FJ Current Flock Size	2	1
Total Current FLock	1.75	1

Matching Covariate	Caliper for IYCF Matching, Households with Children at Baseline	Caliper for IYCF Matching, Households without Children at Baseline
Chicken Income	1.5	1
Egg Income	1.5	0.5
Eggs Eaten	1.5	1.5
WDD Protein	1.5	2

Table 21: Calipers used for IYCF Matching

Caliper values were set so as to produce matched comparisons that were balanced on the matching covariates. If one run of the matching algorithm produced imbalance on a certain input, we tightened the caliper on that variable to ensure better balance. Figure 4 of the main text shows that the matching algorithm and selected calipers greatly improved balance on the matching variables between potential treatment and potential control households. Table 22, below, shows the mean difference and mean standardized difference between treatment and control households.

Matching Variable	Pre-match Difference (T - C)	Post-match Difference (T - C)	Pre-match Standardized Difference (T – C)	Post-match Standardized Difference (T – C)
Distance	0.21	0.03	0.87	0.17
Predicted Value	0.10	0.01	0.75	0.18
Current FJ Owner	0.20	0.00	0.45	0.00
Owned FJ One Year	0.31	0.00	0.43	0.00
FJ Current Flock Size	0.52	0.00	0.28	0.00
Eggs Eaten	1.95	-0.06	0.26	-0.01
Egg Income	7.27	-0.11	0.13	0.00
Chicken Income	32.72	4.76	0.12	0.03
Total Current Flock	1.95	-0.01	0.07	0.00
WDD Protein	-0.07	0.00	-0.14	0.00

Table 22: Balance Table for Direct Effect Match

Tables 23 and 24 show the pre- and post-match balance for the two infant and young child feeding matches: Table 23 for households that had children at baseline (and hence we could match on baseline child nutrition values) and Table 24 for households that did not have children at baseline.

Matching Variable	Pre-match Difference (T - C)	Post-match Difference (T - C)	Pre-match Standardized Difference (T – C)	Post-match Standardized Difference (T – C)
Distance	0.15	0.03	0.65	0.19
Predicted Value	0.06	0.01	0.56	0.22
Child Under 2 Available (at endline)	-0.01	0.00	-0.01	0.00
Child Under 5 Available (at endline)	0.05	0.00	0.18	0.00
Child's Age (at endline)	-0.17	-0.07	-0.15	-0.06
IYCF Egg Consumption	0.06	0.00	0.16	0.00
IYCF Chicken Consumption	0.00	0.00	0.01	0.00
IYCF Protein Consumption	-0.01	0.00	-0.03	0.00
Current FJ Owner	0.15	0.00	0.40	0.00
Owned FJ One Year	-0.06	0.00	-0.24	0.00
FJ Current Flock Size	0.34	-0.01	0.34	-0.05
Total Current FLock	0.32	-0.36	0.06	-0.10
Chicken Income	28.81	4.51	0.15	0.04
Egg Income	3.49	-0.25	0.07	-0.02
Eggs Eaten	0.99	0.46	0.16	0.11
WDD Protein	0.03	0.02	0.07	0.05

Table 23: Balance Table for IYCF Outcomes - Households with Children at Baseline

Matching Variable	Pre-match Difference (T - C)	Post-match Difference (T - C)	Pre-match Standardized Difference (T – C)	Post-match Standardized Difference (T – C)
Distance	0.25	0.02	1.08	0.13
Predicted Value	0.09	0.02	0.59	0.33
Child Under 2 Available (at endline)	-0.05	0.00	-0.11	0.00
Child Under 5 Available (at endline)	-0.12	0.00	-0.24	0.00
Child's Age (at endline)	-0.20	0.05	-0.17	0.05
Current FJ Owner	0.25	0.00	0.54	0.00
Owned FJ One Year	-0.07	0.00	-0.24	0.00
FJ Current Flock Size	0.91	-0.05	0.41	-0.22
Total Current FLock	3.71	-0.15	0.35	-0.05
Chicken Income	-4.46	11	-0.02	0.16
Egg Income	12.05	0.15	0.19	0.01
Eggs Eaten	2.11	0.15	0.20	0.03
WDD Protein	-0.15	-0.03	-0.32	0.05

Table 24: Balance Table for IYCF Outcomes - Households without Children at Baseline