Ankara Üniversitesi
SBF Dergisi,
Erken Görünüm
Türkçe

Araştırma Makalesi / Research Article
DOI

CHICKEN, EGGS, AND REPLICATION *

Emrah Er
Ankara Üniversitesi
ORCID: 0000-0001-9909-7479
ROR ID: ror.org/01wntqw50

● ● ●

Absctract

This study reassesses Thurman and Fisher's (1988) finding on Granger causality between chicken stocks and egg production, focusing on data sources, model selection, and measurement definitions. The U.S. series for 1930–1983 is reconstructed from different editions of USDA *Agricultural Statistics*; 66 alternative source selections are compared with the original results, and the preferred series is extended through 2025. Estimates using the preferred series and logarithmic growth rates support Granger causality from egg production to chicken stocks, but not the reverse. In the extended U.S. sample, the evidence weakens because structural breaks make the estimates reflect an average rather than a stable relationship. Although the FAO panel supports the same direction of causality, Eurostat laying-hen data and the common FAO–Eurostat sample do not consistently confirm it. Reproducibility therefore remains sensitive to stock definitions, data sources, and sample coverage.

Keywords: Replication, Granger causality, Poultry economics, Time series, Panel data

Tavuklar, Yumurtalar ve Replikasyonlar

Öz

Bu çalışma, Thurman ve Fisher'ın (1988) tavuk stokları ile yumurta üretimi arasındaki Granger nedenselliği bulgusunu veri kaynağı, model seçimi ve ölçüm tanımı açısından yeniden değerlendirmektedir. 1930–1983 dönemine ait ABD serisi, USDA *Agricultural Statistics* yıllıklarının farklı basımlarından yeniden oluşturulmuş; 66 alternatif kaynak seçimi özgün sonuçlarla karşılaştırılmış ve tercih edilen seri 2025'e kadar genişletilmiştir. Tercih edilen seri ile tarihsel örnekleme ilişkin logaritmik büyüme oranı tahminleri, yumurta üretiminden tavuk stoklarına doğru Granger nedenselliğini desteklerken ters yönde anlamlı kanıt sunmamaktadır. Genişletilmiş ABD örnekleminde bulgular yapısal kırılmalar nedeniyle zayıflamakta ve dönemler boyunca sabit bir ilişkiden çok ortalama bir ilişkiyi yansıtmaktadır. FAO paneli aynı yönlü nedenselliği desteklese de Eurostat yumurtacı tavuk verileri ve ortak FAO–Eurostat örneklemi bu sonucu istikrarlı biçimde doğrulamamaktadır. Bulgular, özgün sonucun yeniden üretilebilirliğinin stok tanımı, veri kaynağı ve örneklem kapsamına duyarlı olduğunu göstermektedir.

Anahtar Sözcükler: Replikasyon, Granger nedenselliği, Kümes hayvanları iktisadı, Zaman serileri, Panel veri

* Makale geliş tarihi / Received date: 06.05.2026
Makale kabul tarihi / Accepted date: 15.06.2026
Erken görünüm tarihi / Published online: 24.07.2026

Yazar email: eer@politics.ankara.edu.tr

Chickens, eggs, and Replication

Introduction

The old chicken-and-egg question became an econometric result in Thurman and Fisher (1988), hereafter T&F. In annual U.S. data for 1930–1983, T&F find that lagged egg production predicts chicken inventories whereas lagged chicken inventories do not predict egg production. The result is hard to audit today for two reasons. First, T&F cite U.S. Department of Agriculture (USDA) yearbooks only as "various years," without identifying the source vintages or construction rules behind the annual series. Second, their published levels regressions are not checked for nonstationarity or the assumption failures that would weaken the reported Granger tests. This paper therefore reconstructs a documented historical data path, reproduces the published table as closely as possible, and then asks whether the directional result survives modern diagnostics, corrected specifications, and an extension of the U.S. sample beyond 1983. The reconstruction extracts source vintages, enumerates plausible series, scores them against the published result, and tests whether nearby source paths change the conclusion. This makes the data work auditable without claiming to recover the exact series T&F used. The exercise is deliberately conservative. It does not treat replication as a search for a convenient modern data source that happens to produce the same sign pattern. It starts from the publication trail that a reader of the original paper could have followed, preserves the annual units used in the original table, and records where later yearbooks revise earlier calendar years. That choice matters because the data problem is not only missing documentation. It is also a measurement problem created by serial revisions, stock-flow timing, and changes in the poultry sector over nearly a century of observations. The same empirical question can look different when chickens mean laying-capacity inventories, total chicken stocks, or a short European laying-hen panel.

In the reconstructed U.S. data, T&F's directional ordering survives, but their levels specification and interpretation do not. The closest documented reconstruction recovers the published levels pattern. Egg output predicts later chicken stocks, while chicken stocks do not predict later egg output. The raw-level model is treated as a historical replication target, not as the preferred specification. After the specification is corrected, the U.S. growth-rate estimates continue to point from egg-output growth to later chicken-stock growth, not the reverse. The retained equations pass the key serial-correlation diagnostics, and robust Wald inference preserves the main historical ordering. The paper then asks whether the same timing relation appears in international panels built from Food and Agriculture Organization (FAO) data on total chicken stocks and Eurostat data on laying-hen stocks. Those international results are more conditional because FAO and Eurostat measure different chicken stocks on different time supports, and the panel tests average over heterogeneous countries.

That conditionality is plausible because chicken inventories are stock variables, while egg production is a flow. Egg output may lead later inventories when current production changes hatching, retention, or replacement decisions, while inventories may lead later egg output when stocks measure capacity before production appears in the egg series. The contribution is therefore to test how T&F's egg-to-chicken predictive ordering depends on data construction, sample period, and stock definition. This distinction also separates predictive timing from biological priority. Granger evidence is about whether past values of one measured series improve forecasts of another measured series after the latter's own history is controlled (Granger, 1969). It does not identify a structural production equation, and it does not settle the folk question of which biological entity came first. For the poultry data, that narrower interpretation is an advantage. It lets the analysis ask whether producers' and reporting agencies' measured adjustment paths line up in a stable way once the historical source path, stationarity problem, and stock definition are made explicit.

This framing puts the paper closer to agricultural supply response than to a literal origin story. In that literature, observed agricultural quantities adjust gradually because producers form expectations and change durable or biologically constrained production plans over time (Nerlove, 1956, 1979; Rao, 1989). The poultry evidence makes the timing problem concrete. Egg, broiler, and turkey models show how staged production turns current testing, hatching, placement, and supply decisions into later output (Chavas & Johnson, 1981; Chavas & Johnson, 1982). Later broiler studies add that those decisions respond not only to expected prices but also to expected price variance (Aradhyula & Holt, 1989; Holt & Aradhyula, 1990). Work on contracting and vertical coordination then shows why the adjustment path also depends on how growers, integrators, and processors organize production (Crespi & Saitone, 2018; Vukina, 2001). Read together, this literature makes the Granger exercise a test of predictive timing within a measured production system, not a structural claim that egg production mechanically causes the underlying chicken population in every country, period, or stock definition. The empirical question is narrower. Given a documented data path and a stated transformation, does past egg-output growth contain incremental information about later chicken-stock growth?

The closest prior work leaves open the historical data path, the later U.S. sample, and the role of stock definitions across datasets. Patunru (2004) shows why unit roots and cointegration matter for the original exercise, but does not rebuild the historical data across yearbook vintages or test the later U.S. sample. Shikida et al. (2021) takes the question to Brazil and finds bidirectional predictive content. Diaz (2025) extends the poultry timing question to Philippine chickens, eggs, and ducks. Both are useful single-country time-series extensions, but neither reconstructs the original U.S. data path or compares stock concepts across international panels. The additional poultry-supply literature is useful for a different reason. It shows that annual poultry quantities adjust through biological lags, expectations, investment, contracting, and risk rather than through a frictionless one-period response. Demand shifts, differentiated egg demand, welfare-related production systems, and broiler consolidation also show why poultry timing can change across periods and stock definitions (Bakhtavoryan et al., 2021; Neves et al., 2021; Saitone et al., 2025; Thurman, 1987). That setting makes a short annual Granger exercise sensitive to lag choice, deterministic terms, and residual behavior. It also means that a later international comparison should not be expected to reproduce the U.S. historical result mechanically. Different data systems can measure different parts of the production system, and a short laying-hen panel can be closer conceptually while still weaker statistically.

That literature motivates the paper's structure. The Data and Empirical Framework section documents the historical U.S. reconstruction and the later U.S., FAO, and Eurostat samples. The reproduction of the historical U.S. result and the corrected and extended U.S. evidence test whether T&F's ordering survives stationarity, lag-choice, and robustness checks. The international panel evidence tests whether the ordering survives changes in stock definition and country-year support. The conclusion returns to what Granger timing can and cannot say about poultry production.

1. Data and Empirical Framework

T&F estimate bivariate Granger regressions in levels on a short U.S. sample and report the direction of predictive content. The historical series are not documented at the level now expected for replication work, and the levels regressions are not checked for nonstationarity, residual-assumption failures, or robust-inference sensitivity. This paper therefore implements assumption checks and modern time-series and panel diagnostics that were absent from T&F's published analysis. Some address assumptions T&F explicitly relied on, while others are later robustness tools. The empirical design keeps those checks tied to the data problem rather than treating them as a separate exercise.

The historical U.S. data are reconstructed closely enough to reproduce T&F's published table, the historical sample is evaluated with these diagnostics, the U.S. series is extended through 2025, and a long OECD FAO panel built from total chicken stocks is compared with a shorter European-OECD Eurostat panel built from laying hens. The empirical framework is therefore organized around three comparisons rather than one pooled test. The first comparison is forensic. It asks how close a documented yearbook path can get to the printed T&F result. The second comparison is diagnostic. It asks how the answer changes when the same historical window is estimated in stationary growth rates and checked with robust inference. The third comparison is external. It asks whether the timing pattern appears when the U.S. series is extended and when the analysis moves to international data with different stock concepts. Keeping these comparisons separate avoids giving a single table more authority than the data construction can support.

2. Data

The U.S. replication starts from USDA Agricultural Statistics yearbooks (USDA, 2026). For the U.S. chicken and egg series, T&F's source note cites USDA Agricultural Statistics as "U.S. Department of Agriculture, 1983 and others." The note points to multiple yearbook editions, but it does not say which edition supplied each calendar-year value or how revisions across editions were handled. Each yearbook reports a table for a range of past calendar years, and later editions sometimes revise earlier values. The reconstruction in this study treats each edition-year table as a separate source version. For example, the chicken values printed in the 1986 yearbook are one source version, and the values printed in the 1987 yearbook are another. Because the original source note does not identify which source version enters each calendar year, the reconstruction builds 66 candidate U.S. series from different source-selection rules. Each candidate pairs one chicken series with one egg series for 1930–1983. Some candidates use a single edition. Others use priority rules, such as the earliest available value, the latest available value, or the edition closest to the calendar year. The fallback candidates start with a recent yearbook and fill any missing earlier years from progressively older yearbooks. The extension uses the same fallback logic, but gives priority to recent USDA National Agricultural Statistics Service (NASS) QuickStats values when they are available (USDA

NASS, 2026) and then steps backward through the yearbook editions. Appendix A checks the stability of this publication-vintage change. QuickStats does not cover the full historical sample. It begins in 2008 for chicken inventories and in 2013 for egg production. Because the analysis ends in 2025, the final series uses NASS values through 2025 and yearbook values before the NASS window. The resulting U.S. series covers 1930–2025. Appendix A reports the publication-vintage stability check. The reconstructed and extended U.S. series keep T&F's reported variable concepts. Chickens are inventories excluding commercial broilers, and eggs are annual egg output in millions of dozens.

The international extension uses two panels with different chicken-stock concepts. The main panel is drawn from the FAO QCL production database (FAO, 2026) and covers all 38 current OECD members with usable FAO support from 1961 to 2024, subject to country-specific availability. In FAO, chickens are total chicken stocks, including broilers and other birds outside the laying-hen stock most closely tied to egg capacity. FAO eggs are Eggs Primary production. FAO reports this series in tonnes, and the panel data express it in million tonnes for readability rather than converting it to dozens. The panel estimates use log growth rates, so this scale choice does not change within-country growth rates, but the FAO level unit remains different from the U.S. yearbook and NASS egg series. The Eurostat panel is drawn from Eurostat's production of eggs for consumption and number of laying hens table (Eurostat, 2026). It uses laying hens producing eggs for consumption as the chicken stock and eggs for consumption as the egg flow. Compared with FAO, Eurostat has much shorter support. It covers 18 European countries over 2013–2024 in the raw data, and the usable sample is smaller after requiring consecutive annual observations with both chicken and egg values. Eurostat is therefore used only as a laying-hen stock-concept check. The egg series also come from different reporting systems, so the FAO-Eurostat exercise is a construct-validity comparison, not a one-to-one data duplication. Throughout the paper, U.S. chicken inventories mean chickens excluding commercial broilers, FAO chicken stocks mean total chicken stocks, and Eurostat laying-hen stocks mean hens producing eggs for consumption. The distinction is important for interpretation. The U.S. series and Eurostat series are closer to laying capacity, although they differ in time span and reporting system. The FAO series is much longer and broader. It can capture total poultry-sector adjustment, but it also includes birds that are not direct egg-production capacity. For this reason, agreement between FAO and the U.S. historical result is treated as complementary evidence, while disagreement between FAO and Eurostat is treated as information about measurement and support rather than as a simple contradiction. The matched FAO-Eurostat exercise is therefore a sensitivity check, not a decisive horse race between sources.

Table 1 summarizes the main samples used in the paper. It reports coverage, chicken-stock concept, and sample support.

Table 1: Main Samples Used in the Analysis

Sample	Stock Concept	Coverage	Support
Historical U.S.	Inventories excluding commercial broilers	1930–1983	54 annual observations
Extended U.S.	Inventories excluding commercial broilers	1930–2025	96 annual observations
FAO panel	Total chicken stocks, including broilers	38 countries, 1961–2024	2153 country-years
Eurostat panel	Laying hens producing eggs for consumption	18 countries, 2013–2024	152 complete country-years

3. Granger Specifications

The original study can be summarized as a pair of bivariate autoregressions in levels. Let C_t denote chicken stocks and E_t denote egg output. For lag order p, the U.S. system is

$$\begin{aligned} C_t &= a_C^L + \sum_{\ell=1}^{p} \phi_{CC,\ell}^L C_{t-\ell} + \sum_{\ell=1}^{p} \phi_{CE,\ell}^L E_{t-\ell} + u_{Ct}^L, \\ E_t &= a_E^L + \sum_{\ell=1}^{p} \phi_{EE,\ell}^L E_{t-\ell} + \sum_{\ell=1}^{p} \phi_{EC,\ell}^L C_{t-\ell} + u_{Et}^L. \end{aligned} \tag{1}$$

The superscript L marks the historical level system. In $\phi_{CE,\ell}^L$, the first subscript is the dependent variable and the second is the lagged predictor, so $\phi_{CE,\ell}^L$ is the coefficient on lagged eggs in the chicken equation. Eggs Granger-cause chickens if the null hypothesis $H_0\colon \phi_{CE,1}^L = \cdots = \phi_{CE,p}^L = 0$ is rejected, and chickens Granger-cause eggs if $H_0\colon \phi_{EC,1}^L = \cdots = \phi_{EC,p}^L = 0$ is rejected. The historical reconstruction keeps the original structure so the benchmark stays comparable to the original article.

For the corrected U.S. analysis, the same bivariate structure is used in log growth rates. Let $g_t^C = \Delta\log C_t$ and $g_t^E = \Delta\log E_t$. The re-estimated system is

$$
\begin{aligned}
g_t^C &= a_C^G + \sum_{\ell=1}^{p} \phi_{CC,\ell}^G \, g_{t-\ell}^C + \sum_{\ell=1}^{p} \phi_{CE,\ell}^G \, g_{t-\ell}^E + u_{Ct}^G, \\
g_t^E &= a_E^G + \sum_{\ell=1}^{p} \phi_{EE,\ell}^G \, g_{t-\ell}^E + \sum_{\ell=1}^{p} \phi_{EC,\ell}^G \, g_{t-\ell}^C + u_{Et}^G.
\end{aligned} \tag{2}
$$

The superscript G marks the U.S. growth-rate system. The predictive-content restrictions are the growth-rate analogues of the level restrictions. Eggs predict chickens when the joint null on lagged egg-growth coefficients is rejected in the chicken-growth equation, and chickens predict eggs when the joint null on lagged chicken-growth coefficients is rejected in the egg-growth equation.

For country i in year t, the panel specification takes the same form.

$$
\begin{aligned}
g_{it}^C &= a_{Ci}^P + \sum_{\ell=1}^{p} \phi_{CC,\ell}^{P,i} \, g_{i,t-\ell}^C + \sum_{\ell=1}^{p} \phi_{CE,\ell}^{P,i} \, g_{i,t-\ell}^E + u_{Cit}^P, \\
g_{it}^E &= a_{Ei}^P + \sum_{\ell=1}^{p} \phi_{EE,\ell}^{P,i} \, g_{i,t-\ell}^E + \sum_{\ell=1}^{p} \phi_{EC,\ell}^{P,i} \, g_{i,t-\ell}^C + u_{Eit}^P.
\end{aligned} \tag{3}
$$

For the international panels, the analysis uses the Dumitrescu-Hurlin panel Granger test (Dumitrescu & Hurlin, 2012). This test keeps the same maximum lag order across countries while allowing the lag coefficients to vary by country. That flexibility matters because OECD poultry sectors need not share the same adjustment speed. The reported Z-tilde p-values are read as rejections only when the standardized statistic is positive, because the relevant non-causality alternative is an upper-tail average-Wald departure. This rule is applied consistently in the panel interpretation. It prevents a negative standardized statistic with a small two-sided probability from being described as evidence for the reverse Granger direction. That issue matters in the restricted FAO checks, where the sign of the statistic carries the substantive interpretation. The analysis therefore reports the direction of the evidence together with sample support and diagnostic checks, rather than treating statistical significance alone as the conclusion.

4. Results

The results proceed in three steps. First, the original T&F levels exercise is reproduced. Second, the 1930–1983 U.S. specification is corrected by using growth rates, separating the lag-search workflow from T&F's reporting grid, and retaining the original one- through four-lag grid as a comparability check. Third, the same corrected specification is applied to the pooled 1930–2025 U.S. extension before the international panels are examined.

5. Reproducing the Historical U.S. Result

T&F are first reproduced on their own terms using annual U.S. levels, the 1930–1983 window, and the one- through four-lag grid. To choose the historical reconstruction, the original levels specification is estimated for each candidate U.S. series, and the resulting eight-cell table (two directions by four lags) is compared with T&F's published table. The selected data set minimizes the average distance from the published statistics. Table 2 shows that the reconstructed data closely recover their result.[1] The chicken-to-egg tests are far from rejection, with reconstructed p-values from 0,240 to 0,871. The egg-to-chicken tests reject at lags 2, 3, and 4. At lag 2, the closest published comparison, T&F report an F statistic of 10,36. The reconstruction gives 11,63. That 12 percent upward miss on the most diagnostic cell is important. The selected series is the closest documented overall table match, not an exact recovery of every published statistic.[2] Each row reports a directional F test from the relevant equation of an unrestricted bivariate VAR(p) in levels. The p columns are F-test p-values.

1 The public ChickEgg data set distributed with the lmtest package (Hothorn et al., 2022) yields the same qualitative direction but does not match the published statistics as closely as the documented cascade reconstruction. The package documentation says the data were provided by Walter Thurman and made available for R by Roger Koenker, and notes that they are slightly different from the data analyzed in T&F.

2 Appendix A shows how the selected reconstruction sits within the candidate set. All historical replication statistics are conditional on selecting the closest documented source path from that set. The matching score is a ranking heuristic for forensic replication, not an estimator of the true historical data path. The two top cascade candidates tie because they produce the same in-sample values over T&F's 1930-1983 estimable window. The 1986 cascade is therefore a reporting convention rather than a substantive tie-break.

Table 2: Published and Reproduced Historical Granger Statistics

Direction	Lag	*N*	Published *F*	Reproduced *F*	Published *p*	Reproduced *p*	Published R^2	Reproduced R^2
Chicken -> Egg	1	53	0,04	0,03	0,850	0,871	0,96	0,97
	2	52	1,71	1,47	0,190	0,240	0,97	0,97
	3	51	1,10	1,16	0,360	0,338	0,97	0,97
	4	50	0,79	0,85	0,540	0,500	0,97	0,97
Egg -> Chicken	1	53	1,23	1,18	0,270	0,282	0,73	0,72
	2	52	10,36	11,63	<0,001	<0,001	0,81	0,81
	3	51	5,85	6,44	0,0019	0,0010	0,81	0,82
	4	50	4,71	5,06	0,0032	0,0021	0,82	0,82

The reconstructed levels result reproduces T&F's qualitative pattern, but the diagnostics point away from the levels model as the preferred modern estimate.[3] ADF (Dickey & Fuller, 1979; Said & Dickey, 1984) and Phillips-Perron (Phillips & Perron, 1988) tests do not reject the unit-root null in the historical level series, while first differences and growth rates reject the unit-root null and fail to reject stationarity under KPSS (Kwiatkowski et al., 1992). Residual-diagnostic and robust-inference checks add further caution. The original levels regressions are therefore kept for historical comparability. The next section keeps the historical window but changes the specification.

6. Corrected and Extended U.S. Evidence

The corrected U.S. analysis keeps the original question but changes the estimating equation. First, it uses log growth rates rather than levels because the transformed series are stationary and scale-free across the long sample. Second, it separates lag selection from T&F's four-lag reporting grid.[4] The search starts from the information-criterion lags, increases the annual lag length, and then tries trend and BIC-selected break-shift terms. The reported U.S. growth-rate equations pass the key serial-correlation

3 Appendix B reports the historical unit-root and stationarity checks, residual-diagnostic checks, and robust-inference checks for the raw-level equations.

4 Appendix B reports AIC, HQ, BIC, and final prediction error over each sample's feasible diagnostic grid, as well as the workflow search.

checks, and robust Wald inference addresses the remaining heteroskedasticity concerns.[5]

Table 3 reports the corrected historical tests. The table reports *F* tests from unrestricted bivariate log-growth Granger regressions for each direction and lag length. Across the one- through four-lag grid, the stationary growth-rate tests keep T&F's directional ordering. Egg-output growth predicts later chicken-stock growth, while chicken-stock growth does not predict later egg-output growth. Robust Wald inference preserves that result.

Table 3: Corrected Historical U.S. Granger Tests

Direction	Lag	*N*	*F* statistic	*p*-value
Egg growth predicts chicken growth	1	52	14,19	0,001
	2	51	5,83	0,006
	3	50	4,51	0,008
	4	49	4,51	0,004
Chicken growth predicts egg growth	1	52	2,21	0,143
	2	51	1,37	0,265
	3	50	0,90	0,447
	4	49	0,75	0,564

Table 4 shows that the conventional eggs-to-chickens ordering also appears in the 1930–2025 U.S. sample. The reverse direction is essentially absent throughout the one- through four-lag grid. The full-sample robust Wald evidence is strongest in the one-lag equation, while longer-lag estimates are reported as comparison estimates. Because the extension spans several poultry-market regimes, the final check asks whether the full-sample coefficients are stable over time.[6] The full-sample one-lag equations reject stability, with BIC-selected breakpoints in 1965 for the egg-to-chicken equation and 1969 for the chicken-to-egg equation. These dates are estimated coefficient breaks, not separate evidence of a new Granger direction. They imply that the 1930–2025 coefficients

5 The reported U.S. growth-rate specifications do not show residual serial correlation. Appendix B reports the remaining residual diagnostics and heteroskedasticity-consistent Wald checks (MacKinnon & White, 1985; White, 1980).

6 Appendix B reports the unknown-break checks for the one-lag equations used in the structural-break analysis.

summarize changing regimes rather than a time-invariant predictive relation. The corrected U.S. exercise therefore supports a more specific claim than the original levels result. Once levels are abandoned, the growth-rate estimates point from egg-output growth to later chicken-stock growth.

Table 4: Extended U.S. Granger Tests

Direction	Lag	*N*	*F* statistic	*p*-value
Egg growth predicts chicken growth	1	94	14,13	0,001
	2	93	5,21	0,007
	3	92	4,05	0,010
	4	91	4,17	0,004
Chicken growth predicts egg growth	1	94	0,00	0,971
	2	93	0,09	0,917
	3	92	0,13	0,944
	4	91	0,17	0,954

7. International Panel Evidence

The panel analysis follows the same corrected logic as the U.S. extension. It uses growth rates first, then asks whether the direction survives support and dependence checks. FAO and Eurostat answer related but not identical measurement questions, so each panel is first estimated on its own support. The matched-support exercise is a sensitivity check, since differences between the two panels reflect time depth, country coverage, and stock definition at the same time. Because countries can share year-specific movements in feed costs, disease conditions, demand, or reporting practice, year-demeaned and dependence-aware checks are reported before the panel Granger results are interpreted.

Table 5 reports the baseline panel Granger tests.[7] The FAO statistic is one-directional. Egg-output growth predicts total-chicken-stock growth at lags 1 through 3, while the reverse direction is not rejected. Because FAO measures

7 The panel unit-root diagnostics support using growth rates, with stronger support in FAO than in Eurostat. Appendix C reports the panel unit-root diagnostics. In FAO, the level evidence is mixed, while growth rates reject the unit-root null for both series under both IPS and CIPS tests (Im et al., 2003; Pesaran, 2007). Eurostat points in the same direction under IPS, but its shorter support rules out the CIPS check.

total chicken stocks, the result concerns broader poultry-stock dynamics rather than the U.S. laying-capacity concept.

Table 5: Panel Granger Results

Dataset	Direction	Lag	*N*	Statistic	*p*-value
FAO	Egg -> Chicken	1	2.012	7,22	<0,001
		2	2.012	7,43	<0,001
		3	2.012	5,28	<0,001
	Chicken -> Egg	1	2.012	-1,12	0,261
		2	2.012	1,90	0,057
		3	2.012	1,32	0,186
Eurostat	Egg -> Chicken	1	105	0,66	0,508
	Chicken -> Egg	1	105	1,17	0,241

The same direction appears after simple checks for common year movements.[8]

The shorter Eurostat laying-hen panel is less decisive. Its baseline test rejects neither direction, and the adjusted checks do not give a stable alternative ordering. Eurostat is therefore a useful stock-concept check, but it does not overturn the FAO evidence. For Eurostat to play the same role as FAO, it would need enough consecutive country-year support to estimate the same annual lag structure without most of the identifying variation coming from a short post-2013 window.

Sensitivity checks narrow that interpretation rather than erase it.[9] The FAO result is not a claim that most country-level tests reject. Restricting FAO to Eurostat countries preserves the egg-to-chicken rejection. The reverse statistic is negative, so the upper-tail Granger rejection is absent despite the small two-sided Z-tilde p-value. The pattern is not recovered when FAO is restricted to Eurostat years or to the fully matched overlap. Thus the international evidence is strongest as a long-panel total-chicken-stock result, not as a stock-concept-invariant claim. This is the main reason the international exercise is presented as a robustness and

8 Appendix C reports the year-demeaned panel tests, the fixed-effects check with Driscoll-Kraay standard errors (Driscoll & Kraay, 1998), and the fixed-effect equation diagnostics. The Driscoll-Kraay fixed-effect check keeps the FAO result one-directional while allowing for cross-sectional dependence in the covariance matrix.

9 Appendix C reports country-level heterogeneity, restricted-FAO, matched-overlap, and stock-ratio checks.

interpretation step rather than as a replacement for the U.S. reconstruction. The FAO panel has enough time depth to estimate annual dynamics across many countries, but it measures a wider poultry stock. The Eurostat panel measures a stock closer to laying capacity, but its usable support is short and uneven. The matched overlap gives the cleanest comparison of definitions, yet it also removes much of the time variation that makes Granger tests informative. The empirical message is therefore deliberately bounded. The long FAO panel is consistent with the corrected U.S. timing result, while the laying-hen robustness checks do not show that the direction is invariant to stock definition and sample support.

Conclusion

The evidence supports a narrower answer to the chicken-and-egg question than the published T&F result suggests. A documented reconstruction of the 1930–1983 U.S. series reproduces their qualitative levels result. Egg output predicts later chicken stocks, while chicken stocks do not predict later egg output. Once the nonstationarity evidence is taken seriously, however, the better-supported transformation is log growth rates. In that corrected historical sample, conventional and robust tests point from egg-output growth to later chicken-stock growth. The pooled U.S. extension through 2025 has the same conventional ordering. Robust inference is strongest in the one-lag equation, and the broader lag grid still points away from chicken-stock growth predicting later egg-output growth. Structural-break checks reject coefficient stability in the full-sample one-lag equations, with BIC-selected breakpoints in 1965 and 1969. The extended U.S. evidence is therefore best summarized as an average relation across changing regimes.

The international panels narrow the stock-concept claim rather than erase the timing result. The long FAO panel gives one-directional eggs-to-total-chickens evidence in the baseline table and in dependence-aware checks. The shorter Eurostat laying-hen panel is less decisive. The FAO country restriction preserves egg-to-chicken evidence but does not produce an upper-tail reverse-direction rejection. The pattern does not survive restriction to Eurostat years or the fully matched overlap. The panel evidence therefore supports broader poultry-stock timing in the long FAO data, while it does not establish a stock-concept-invariant result.

The contribution is to turn T&F’s undocumented source note into an auditable replication and extension. The paper reconstructs candidate historical series, documents the selected data path, applies time-series checks absent from the original paper, and compares modern U.S. and international evidence. The result is not a structural solution to the chicken-and-egg riddle. It is a reproducible measurement and predictive-timing exercise whose strongest evidence remains the corrected U.S. series. The long FAO total-chicken-stock panel adds complementary international evidence with a broader stock definition.

The paper therefore leaves a clear empirical target for future work. Longer laying-hen panels would show whether the egg-to-chicken timing found in the reconstructed U.S. data is a feature of laying capacity itself or partly a feature of the broader poultry stock measured in FAO. The places where the result weakens are also informative. They identify where source support, stock definitions, and structural change limit the reach of a simple Granger-causality table

For ethical issues: sbfdergi@politics.ankara.edu.tr

Plagiarism check: Completed.

Declaration of conflicting interests: The author(s) declare no conflict of interest.

References

Aradhyula, S. V., & Holt, M. T. (1989). Risk behavior and rational expectations in the U.S. broiler market. *American Journal of Agricultural Economics, 71*(4), 892–902. https://doi.org/10.2307/1242667

Bakhtavoryan, R., Hovhannisyan, V., & Devadoss, S. (2021). An empirical evaluation of egg demand in the United States. *Journal of Agricultural and Applied Economics, 53*(2), 280–300. https://doi.org/10.1017/aae.2021.9

Chavas, J.-P., & Johnson, S. R. (1981). An econometric model of the US egg industry. *Applied Economics, 13*(3), 321–335. https://doi.org/10.1080/00036848100000003

Chavas, J.-P., & Johnson, S. R. (1982). Supply dynamics: The case of U.S. broilers and turkeys. *American Journal of Agricultural Economics, 64*(3), 558–564. https://doi.org/10.2307/1240650

Crespi, J. M., & Saitone, T. L. (2018). Are cattle markets the last frontier? Vertical coordination in animal-based procurement markets. *Annual Review of Resource Economics, 10*(1), 207–227. https://doi.org/10.1146/annurev-resource-100517-022948

Diaz, J. F. T. (2025). Which came first, the chicken or the egg? What about ducks?: Granger causality using Philippine poultry data. *Journal of Economic Statistics, 3*(1), 1–6. https://doi.org/10.58567/jes03010001

Dickey, D. A., & Fuller, W. A. (1979). *Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427–431*. https://doi.org/10.1080/01621459.1979.10482531

Driscoll, J. C., & Kraay, A. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics, 80*(4), 549–560. https://doi.org/10.1162/003465398557825

Dumitrescu, E.-I., & Hurlin, C. (2012). Testing for Granger non-causality in heterogeneous panels. *Economic Modelling, 29*(4), 1450–1460. https://doi.org/10.1016/j.econmod.2012.02.014

Eurostat. (2026). *Production of eggs for consumption and number of laying hens (apro_ec_egghen)* [Dataset]. https://ec.europa.eu/eurostat/cache/metadata/en/apro_ec_egghen_esmseg.htm

FAO. (2026). *FAOSTAT: Crops and livestock products and livestock stocks* [Dataset]. https://www.fao.org/faostat/en/#data/QCL

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424–438. https://doi.org/10.2307/1912791

Holt, M. T., & Aradhyula, S. V. (1990). Price risk in supply equations: An application of GARCH time-series models to the U.S. broiler market. *Southern Economic Journal, 57*(1), 230–242. https://doi.org/10.2307/1060492

Hothorn, T., Zeileis, A., Farebrother, R. W., Cummins, C., Millo, G., & Mitchell, D. (2022). *lmtest: Testing linear regression models (Version 0.9-40)* [R package]. https://doi.org/10.32614/CRAN.package.lmtest

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics, 115*(1), 53–74. https://doi.org/10.1016/S0304-4076(03)00092-7

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics, 54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics, 29*(3), 305–325. https://doi.org/10.1016/0304-4076(85)90158-7

Nerlove, M. (1956). Estimates of the elasticities of supply of selected agricultural commodities. *Journal of Farm Economics, 38*(2), 496–509. https://doi.org/10.2307/1234389

Nerlove, M. (1979). The dynamics of supply: Retrospect and prospect. *American Journal of Agricultural Economics, 61*(5), 874–888. https://doi.org/10.2307/3180340

Neves, M. F., Gray, A. W., & Lourenço, C. E. (2021). Mantiqueira: Innovating and disrupting in the egg business. *International Food and Agribusiness Management Review, 24*(1), 138–161. https://doi.org/10.22434/ifamr2020.0031

Patunru, A. A. (2004). *Chickens vs. eggs: Replicating Thurman and Fisher (1988)*. https://web.pdx.edu/~crkl/ec571/eggs.pdf

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics, 22*(2), 265–312. https://doi.org/10.1002/jae.951

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika, 75*(2), 335–346. https://doi.org/10.1093/biomet/75.2.335

Rao, J. M. (1989). Agricultural supply response: A survey. *Agricultural Economics, 3*(1), 1–22. https://doi.org/10.1016/0169-5150(89)90036-4

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika, 71*(3), 599–607. https://doi.org/10.1093/biomet/71.3.599

Saitone, T. L., Schaefer, A., Scheitrum, D., Arita, S., Breneman, V., & Nemec Boehm, R. (2025). Consolidation, productivity, and downstream prices in the U.S. poultry industry. *Agricultural and Resource Economics Review, 54*, 157–178. https://doi.org/10.1017/age.2025.15

Shikida, C. D., de Araujo, A. F., Jr., & Figueiredo, E. A. (2021). Which came first? Eggs and chicken in Brazil: A new study. *REGEN, 2*(2), 87–99.

Thurman, W. N. (1987). The poultry market: Demand stability and industry structure. *American Journal of Agricultural Economics, 69*(1), 30–37. https://doi.org/10.2307/1241303

Thurman, W. N., & Fisher, M. E. (1988). *Chickens, eggs, and causality, or which came first? American Journal of Agricultural Economics, 70*(2), 237–238. https://doi.org/10.2307/1242062

USDA. (2026). *Agricultural statistics yearbooks* [Dataset and historical statistical reports]. https://www.nass.usda.gov/Publications/Ag_Statistics/

USDA NASS. (2026). *QuickStats database* [Dataset]. https://quickstats.nass.usda.gov/

Vukina, T. (2001). Vertical integration and contracting in the U.S. poultry sector. *Journal of Food Distribution Research, 32*(2), 29–38. https://doi.org/10.22004/ag.econ.27819

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838. https://doi.org/10.2307/1912934