RIETI Discussion Paper Series 21-E-047

Culture, Tastes, and Market Integration: Testing the localized tastes hypothesis (Revised)

MORI, Tomoya
RIETI

WRONA, Jens
University Duisburg-Essen

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 21-E-047
First Draft: June 2021
Revised: May 2025

Culture, Tastes, and Market Integration: Testing the localized tastes hypothesis*

Tomoya Mori
Institute of Economic Research, Kyoto University
Research Institute of Economy, Trade and Industry

Jens Wrona
University Duisburg-Essen, Mercator School of Management (MSM), Institute of East Asian Studies (IN-EAST)

Abstract

To test the localized tastes hypothesis, we use historical dialect similarity as an instrument to predict the persistent component of regional taste differences. Analyzing wholesale markets for fruits and vegetables in Japan, we find that predicted taste differences have a strong, statistically significant effect, explaining approximately 9% of the mean volatility in law-of-one-price deviations. Our findings are robust across extensive validity checks, which scrutinize and relax our exclusion restriction, distinguishing between various sources of endogeneity, and confirm our baseline results based on alternative instruments, which exploit exogenous differences in agro-climatic endowments.

Keywords: Market Integration, Tastes, Culture, Dialects
JEL classification: D12, F15, N75, Q11, R22, Z13

*This research was conducted as part of the project "Agglomeration-based framework for empirical and policy analyses of regional economies," undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The draft of this paper was presented at the RIETI DP seminar for the paper. We thank the participants of the seminar for their helpful comments. Tomoya Mori acknowledges funding from the Kajima Foundation, the Murata Science Foundation, the International Joint Research Center of Advanced Economic Theory of the Institute of Economic Research in Japan, and the MEXT Japan Grant in Aid for Research No. 21H00706. Jens Wrona acknowledges funding from the German Research Foundation (DFG) as member of the Research Training Group (RTG) 2484: "Regional Disparities and Economic Policy" under Project No. 397383192.

1 Introduction

This study provides robust causal evidence in support of the localized tastes hypothesis, according to which markets with similar historically determined tastes appear to be better integrated. Focusing on Japan's wholesale markets for fruits and vegetables, we show that persistent bilateral commodity-specific taste differences – predicted by the overlap in historical, regional dialects – have a strong, statistically significant effect, accounting for approximately 9% of the mean volatility in law-of-one-price (LOP) deviations at the commodity level.

Distance and border-effect estimates from gravity-type estimations which aim to explain bilateral trade volumes or the volatility of LOP deviations are too large and too persistent to solely reflect the effects of geography.[1] To rationalize these findings, Grossman (1998) once proposed three key explanations: (*i.*) information frictions, that rapidly increase in distance, (*ii.*) network effects and (*iii.*) localized tastes, which are historically determined and persist across time (see also Head and Mayer, 2013, p. 1216). While there is causal evidence on the importance of information frictions (cf. Steinwender, 2018) and the role of networks (cf. Cohen et al., 2017; Parsons and Vézina, 2018), this paper's purpose is to provide complementary causal evidence supporting the localized tastes hypothesis.[2]

In their survey, Head and Mayer (2013) synthesize the localized tastes hypothesis, arguing that before the advent of cheap, refrigerated transportation, regional tastes historically emerged in response to the abundant supply of local varieties (cf. Atkin, 2013). In regions where agro-climatic conditions favored certain crops, localized preferences naturally developed and were passed down through generations (cf. Bisin and Verdier, 2001). Conversely, no such tastes formed in areas where these crops were historically unavailable. Localized tastes that emerged from this process would spread in space through cultural exchange whenever regions were linked by economic interactions such as inter-regional migration (cf. Bronnenberg et al., 2012; Atkin, 2016). Crucially, the regional distribution of tastes that ultimately emerged from the complex interplay between heterogeneous agro-climatic endowments and past cultural exchange would tend to endure, even after their original determinants – such as prohibitive transportation costs – have lost their importance.[3]

[1] The trade literature has documented large trade-reducing effects of international borders (cf. McCallum, 1995; Anderson and van Wincoop, 2003; Chen, 2004) and intra-national borders (cf. Wolf, 2000; Hillberry and Hummels, 2003; Combes et al., 2005; Millimet and Osang, 2007; Yilmazkuday, 2012). Evidence on the persistent effects of historical borders comes from Nitsch and Wolf (2013) and Felbermayr and Gröschl (2014). The magnitude and the persistence of the distance effect on the volume of bilateral trade are discussed among others by Disdier and Head (2008), Yotov (2012), Head and Mayer (2014). See Engel and Rogers (1996), Parsley and Wei (2001) and Ceglowski (2003) for similar results with respect to the effect of inter- and intra-national borders on the volatility of LOP deviations. Parsley and Wei (1996), Crucini et al. (2010, 2015) and Elberg (2016) provide complementary evidence on the effect of distance on the volatility of LOP deviations.

[2] See also Rauch (1999), Huang (2007), Hortaçsu et al. (2009) Lendle et al. (2016) and Bailey et al. (2021) for indirect evidence on the information channel as well as Rauch and Trindade (2002), Herander and Saavedra (2005) and Dunlevy (2006) on the effects of (migrant) networks.

[3] See Colson-Sihra et al. (2023) for anecdotal evidence from France, highlighting a persistent north-west versus

Building on Head and Mayer's (2013) argumentation, we propose to instrument contemporaneous taste differences between Japanese wholesale markets for fruits and vegetables by historical dialect similarity (cf. Falck et al., 2012), which is a comprehensive measure for past cultural exchange, capturing the cumulative impact of all previous interactions that left lasting imprints on local dialects. By drawing on historical dialect maps from the Linguistic Atlas of Japan (cf. National Language Research Institute, 1966; Kumagai, 2016), we are able to map Japan's historical dialect landscape from the beginning of the 20$^{\text{th}}$ century into a robust measure of past cultural exchange, which is highly correlated with the bilateral distance in contemporaneous taste estimates.[4] We contend that this correlation stems from past interactions that, through a process of cultural co-evolution, shaped both the persistent distribution of historical tastes (cf. Bisin and Verdier, 2001) and the development of historical dialects (cf. Hinskens et al., 2005; Schmidt, 2010). The similarity of historical dialects across Japanese regions therefore can be used to predict the persistent component of the contemporaneous preference differences, which ultimately enables us to identify the causal effect of taste heterogeneity on the integration of Japan's wholesale markets for fruits and vegetables.

In order to assess the bilateral integration of 72 Japanese wholesale markets for fruits and vegetables, we follow Engel and Rogers (1996) and Crucini et al. (2010), who hypothesize that the price volatility of similar goods sold in different locations is related to a set of gravity variables separating these markets. Measuring the deviation from the law of one price for a given commodity and a given market pair by the log of the relative market price, we are adopting the time series volatility of the LOP deviation as our referred measure for market integration. According to Head and Mayer (2021, p. 39) the volatility of LOP deviations is a "natural complement" to measuring the integration of markets based on bilateral trade volumes, which in a setting like ours is impractical because most markets are not integrated through direct arbitrage but indirectly linked through a networks of shared regional suppliers. Using Japan's Fresh Fruits and Vegetables Wholesale Market Survey, which tracks the prices of 31 fruits and vegetables across 72 wholesale markets on a monthly basis from 2010 to 2021, we not only confirm the positive and highly significant effect of geographic distance on the volatility of LOP deviations (cf. Parsley and Wei, 1996; Engel and Rogers, 1996), but also show that this effect is reduced by about 40% when accounting for the structure of supply networks.

Following Colson-Sihra et al. (2023), regional preference differences are measured as the absolute differences of market-specific structural taste estimates for 31 fruits and vegetables, which

south-east divide in the consumption patterns for butter versus olive oil.

[4]The Linguistic Atlas of Japan is the most comprehensive data source on the spatial distribution of historical dialects in Japan, consisting of 300 dialect maps on the dialectal forms and pronunciations for 240 survey items from approximately 2,400 locations across Japan.

we obtain from estimating an Almost Ideal Demand System (AIDS) (cf. Deaton and Muellbauer, 1980; Atkin, 2013). For the taste estimation we rely on the 2019 wave of the Japanese Family Income and Expenditure Survey, which provides micro-level data on the food expenditures and the socio-economic characteristics of Japanese households at the spatially disaggregated level of 168 municipalities. Due to the additive separability of the AIDS, commodity-specific taste parameters can be identified as municipality-specific components of food expenditure shares that can not be explained by the price vector, total food expenditure, or household characteristics.

In order to test the localized tastes hypothesis, we conduct a simple two-stage-least-squares estimation: In the first stage, we use the similarity of historical dialects across regions as a proxy for past cultural exchange to isolate the persistent variation in bilateral taste dissimilarity, which in the second stage is used to identify the causal effect of regional taste heterogeneity on the integration of regional wholesale markets, as reflected by the volatility of LOP deviations. The point estimate of the bilateral taste distance – instrumented with historical dialect similarity – has the expected positive sign, is statistically significant from zero, and explains about 9% of the average volatility of LOP deviations in our sample. Moreover, our first-stage results show that historical dialect similarity has a strong negative effect on bilateral taste distance, which not only lends credibility to our instrumental variable approach but also is consistent with the predictions of the localized tastes hypothesis (cf. Head and Mayer, 2013).

The Japanese setting in many ways is ideal for testing the localized tastes hypothesis. Few high-income countries devote such a significant portion of domestic consumption expenditure to food, underscoring its cultural importance.[5] Japan's traditional dietary culture (*washoku*) is renowned for its deep historical and regional roots and recognized since 2013 as an Intangible Cultural Heritage of Humanity (cf. UNESCO, 2013).[6] Finally, it is worth noting that neither international trade nor immigration from abroad – which Head and Mayer (2013, p. 1223) identify as potential forces that can disrupt the intra-national taste distribution – play a significant role in Japan.[7]

To identify regional tastes, we focus on Japanese foods, particularly fruits and vegetables, which are arguably relatively stable over time (cf. Colson-Sihra et al., 2023). Covering a significant share of (native) crops that have been consumed in Japan for centuries, allows us to use historical dialect variation to explain the spatial distribution of localized tastes. At the same time, we exploit Japan's unexpected opening to the rest of the world after more than 200 years of isolation (cf. Bernhofen

[5]According to the UN's (2020) National Accounts Statistics Japan's share of total domestic consumption expenditure devoted to food and non-alcoholic beverages in 2019 was with 15.5% higher than for Germany (10.8%), France (13.1%), Italy (14.3%), Spain (12.5%) South Korea (11.4%), the UK (9, 5%) and the US (6.2%).

[6]The Ministry of Agriculture, Forestry and Fisheries (MAFF) has put together a comprehensive list of Traditional Foods in Japan (cf. MAFF, 2025a), highlighting the regional origins and the historical significance of the Japanese cuisine.

[7]Japan has a relatively small share of immigrant workers (cf. OECD, 2024) and consistently ranks low in trade openness (cf. Beltramello et al., 2011). See also Lawrence (1987, 1991) and Saxonhouse (1993) for earlier discussions on Japan's exceptionally low export/import-to-GDP ratio.

and Brown, 2004, 2005) as a natural experiment, that allows us to identify apples as subsample of commodities, that was rather lately introduced to Japan and therefore can be used to check the plausibility of our exclusion restriction. Our exclusion restriction requires that historical dialect similarity does not affect the volatility of LOP deviations through any other channel than through shaping the spatial distribution of tastes. To provide evidence against alternative channels – such as migration and trade networks (cf. Falck et al., 2012; Bauernschuster et al., 2014; Lameli et al., 2015) – that could link historical dialect similarity to the volatility of LOP deviations, we perform a zero-first-stage (ZFS) test, following Altonji et al. (2005), Nunn (2008) and Angrist et al. (2010). The ZFS test is essentially a placebo regression of the instrument on the outcome variable, using a subsample for which it can be plausibly argued that the instrument's first-stage effect on the endogenous explanatory variable is zero.

Based on anecdotal and descriptive evidence on the late introduction and slow popularization of apples in Japan during the 20th century, we argue that regional apple preferences emerged long after the historical dialects documented by the Linguistic Atlas of Japan (LAJ), which is informative about the evolution of historical dialects until the beginning of the 20th century. Reassuringly, we find that a ZFS test on the apple subsample yields tightly estimated near-zero coefficients for the effect of historical dialect similarity on the volatility of LOP deviations, which across all but one specification are statistically indistinguishable from zero. Following van Kippersluis and Rietveld (2018), we use the ZFS test results as an input for Conley et al.'s (2012) local-to-zero (LTZ) approach, which enables robust inference of instrumental variable estimates when the instrument is only plausibly (and not strictly) exogenous. The resulting LTZ estimates are somewhat smaller than our baseline results but remain comparable in magnitude and significance, leading us to conclude that our instrument can be regarded as plausibly exogenous.

Further support for our exclusion restriction comes from a more narrowly defined version of our historical dialect instrument. For the compilation of LAJ map #179, respondents were asked whether they associate the Japanese word for "tuber" or "potato" ("*imo*") with potato (*jagaimo*), sweet potato (*satsumaimo*) or taro (*satoimo*). Using the geographic variation in the responses to this question allows us to construct a narrow instrument for the market-pair-specific taste distances based on a smaller subsample of these three vegetables. The resulting 2SLS estimates for the effect of taste heterogeneity on the volatility of LOP deviations closely align with our baseline results, while the OLS estimates are somewhat smaller but statistically larger than zero. Since the responses from LAJ map #179 are unlikely to be correlated with other factors potentially influencing the relative price volatility of potatos, sweet potatos and taro, we interpret the confirmation of our baseline results as additional supportive evidence in favor of our exclusion restriction.

In addition to Japan's rich history, we also exploit marked differences in geography and climate as sources of exogenous variation, which allow us confirm our baseline IV results based on a set of alternative instruments.[8] Before the advent of cheap mass transportation, consumers were likely to develope localized food preferences for those varieties that were both inexpensive and readily available under the given local agro-climatic conditions (cf. Atkin, 2013). Building on this insight, we use differences in climatic conditions and soil quality to construct a complementary instrument for commodity-specific taste differences between markets. Accounting for the effects of climate and soil quality does not alter the magnitude or significance of the 1$^{\text{st}}$-stage effect of the historical dialect instrument, whose joint validity with the newly added instruments cannot be rejected at conventional significance levels. Reassuringly, the effect of regional taste heterogeneity on relative price volatility remains consistently positive and significant, regardless of which combination of instrumental variables is used for causal identification.

Compared to our preferred 2SLS approach, ordinary least squares (OLS) produces a much smaller point estimates that are statistically indistinguishable from zero. We argue that attenuation bias due to measurement error offers a plausible explanation for this result, because market-specific taste parameters cannot be directly observed and therefore need to be estimated. Using these estimates as proxies for unobservable regional tastes inherently introduces some margin of error, necessitating the use of an instrumental variable approach (cf. Aw et al., 2023).

To rule out omitted variable bias as a possible alternative explanation for the difference between our OLS and 2SLS results, we enhance our baseline specification – which already controls for market- and commodity-specific unobserved heterogeneity through a set of appropriately specified fixed effects – by introducing an even more demanding bilateral fixed effect structure. Adding $4+(4\times 3)/2 = 10$ island-pair-specific fixed effects and $8+(8\times 7)/2 = 36$ region-pair-specific fixed effects to account for all bilateral variation between Japan's 4 main islands and its 8 administrative regions neither alters the magnitude nor affects the significance of the taste effect. However, incorporating a large number of prefecture-pair fixed effects to absorb all bilateral variation between Japan's 46 prefectures introduces a weak-instrument problem, which results in inflated 2$^{\text{nd}}$-stage standard errors. Reassuringly, we find that our baseline results are confirmed, when relying on a weak-instrument-robust inference (cf. Anderson and Rubin, 1949; Stock and Wright, 2000).

The remainder of the paper is organized as follows: Section 2 introduces the data used to measure the integration of fruit and vegetable markets in Section 3 and to estimate localized tastes in Section 4. Section 5 outlines our identification strategy and presents the main results, followed by a robustness analysis in Section 6. Section 7 concludes.

[8] Japan stretches for approximately 2,400 kilometers through the western North Pacific Ocean and has six principal climate zones, which significantly differ in terms of mean temperatures and precipitation levels (cf. JCDP, 2025).

2 Data

2.1 Wholesale prices

The data source for wholesale prices of fruits and vegetables is the Fresh Fruits and Vegetables Wholesale Market Survey conducted by the Ministry of Agriculture, Forestry and Fisheries (MAFF). This survey provides monthly information on the sales values and quantities of 17 fruits and 48 vegetables (see Table A.1, delegated to the Appendix) sold at the 72 wholesale markets in Japan (see Table A.2 and Figure A.1, delegated to the Appendix) over the period from 2010 to 2021.[9] We exclude all transactions involving imports and focus exclusively on domestically produced commodities. The remaining transactions are categorized by their source, either another wholesale market or one of Japan's 46 prefectures (excluding Okinawa). In order to obtain for each market a series of monthly commodity prices, we compute the average market price for this commodity based on all transactions within a given month.[10]

2.2 Family Income and Expenditure Survey

In order to estimate commodity- and location-specific preference parameters that can be used to measure the taste distance between two markets, we use the 2019 wave of the Japanese Family Income and Expenditure Survey conducted by the Ministry of Internal Affairs and Communications (MIC). From more than 700 expenditure categories, we identify 10 fruits and 21 vegetables, that we are able to match to the Fresh Fruits and Vegetables Wholesale Market Survey classifying them into four broad commodity categories: "fruit", "vegetable (root)", "vegetable (leaf)" and "vegetable (fruit)" (see Table A.1, delegated to the Appedix).

The Family Income and Expenditure Survey covers 168 municipalities.[11] In order to combine the Fresh Fruits and Vegetables Wholesale Market Survey with the Family Income and Expenditure Survey, we match the location of 66 wholesale market to their respective municipality and associated the remaining 5 wholesale markets whose municipality is not covered by the Family Income and Expenditure Survey with the geographically closest municipality from the Family Income and Expenditure Survey (see Table A.2, delegated to the Appendix). Following a stratified

[9] Although it is possible to observe market prices on a daily basis (cf. Kano et al., 2013), we decided to work with a monthly frequency. Focusing on monthly data not only gives us a broader coverage of markets, but also allows us to consistently compute relative price series. Daily data, by contrast, would be challenging to standardize due to variations in market opening days.

[10] Missing market prices for commodities subject to seasonal supply variations (e.g. strawberries) are imputed using the most recent available monthly price at the respective market.

[11] The 168 sampled municipalities consist of 52 prefecture capitals and government-designated cities, 74 municipalities with at least 50,000 inhabitants, which are chosen based on the municipality type (urban/rural), population density, population change, industrial structure, and the age composition of the household heads, and 42 municipalities with less than 50,000 inhabitants, which are chosen based on the regional composition, topographical characteristics, and the age composition of household heads.

sampling approach the sampled municipalities are subdivided into unit blocks that roughly share the same population. Between 1300 and 1400 of these blocks are surveyed each month.[12] From each of these blocks six randomly chosen households are surveyed for 6 consecutive months before being replaced by other randomly selected households from the same block.

Because families can be assumed to be more firmly tied to the regions they currently living in and therefore more habituated to localized taste patterns, we are focusing in our analysis on multi-member households.[13] Alongside the household's food expenditure, we incorporate the following household characteristics into our analysis: household income, household size, the number of members below 18 and above 64 years old, as well as the industry (18 categories) and occupation (12 categories) in which the household head is employed.

2.3 The Linguistic Atlas of Japan

The Linguistic Atlas of Japan, published in six volumes between 1966 and 1974, features 300 dialect maps that illustrate the geographic distribution of dialectal forms and pronunciations for 240 survey items collected from 2,400 localities across Japan, which are separated by an average distance of about 12 kilometers (cf. National Language Research Institute, 1966).[14,15] The survey underlying the LAJ was conducted between 1957 and 1965, targeting predominantly (99.7%) male informants, who had lived continuously in the surveyed locality from age 3 to 15 and were mostly (97%) born between 1879 and 1903 (cf. Kumagai, 2016). For our study, we rely on the Linguistic Atlas of Japan Database (LAJDB), which contains coded data corresponding to the published dialect maps for 141 prototypical language characteristics, corresponding to 58% of the 240 questionnaire items that were published as dialect maps covered by the LAJ.[16] A complete list of all 141 survey items with the corresponding LAJ map number covered by the LAJDB is reported in Table A.3 (delegated to the Online Appendix).

Although not the oldest linguistic atlas on Japanese, the LAJ is the most comprehensive data source on the spatial distribution of historical dialects in Japan, and was specifically designed to explore the geographic distribution of standard Japanese as well as the history of individual

[12] Per month 1/12 of the surveyed blocks are randomly replaced by other blocks from the same municipality.

[13] The following household types are excluded from the analysis: student households, households living in hospitals or similar facilities, households living in dwellings with shops, hotels, etc., households with multiple families, households with more than three live-in maids, households whose heads are absent for more than three months, foreign households.

[14] The LAJ's 300 dialect maps have been digitized and made available (`https://mmsrv.ninjal.ac.jp/laj_map/`) by the National Institute for Japanese Language and Linguistics (NINJAL).

[15] In total 285 questionnaire items were surveyed. The number of surveyed localities differs across these 285 questionnaire items, and can be broken down as follows: 128 items were surveyed at approx. 2,400 localities, 36 items were surveyed at approx. 2,000 survey localities, 55 items were surveyed at approx. 1,700 localities, 62 items were surveyed at approx. 1,000 localities, and 4 items were surveyed at approx. 400 localities (cf. National Language Research Institute, 1966).

[16] See Kumagai (2016) for details on the construction of the Linguistic Atlas of Japan Database (LAJDB).

words (cf. National Language Research Institute, 1966).[17] Figure 1 reproduces LAJ Map #262 (*icicle*) as an illustrative example for the 300 detailed dialect maps included in the LAJ.[18] Each

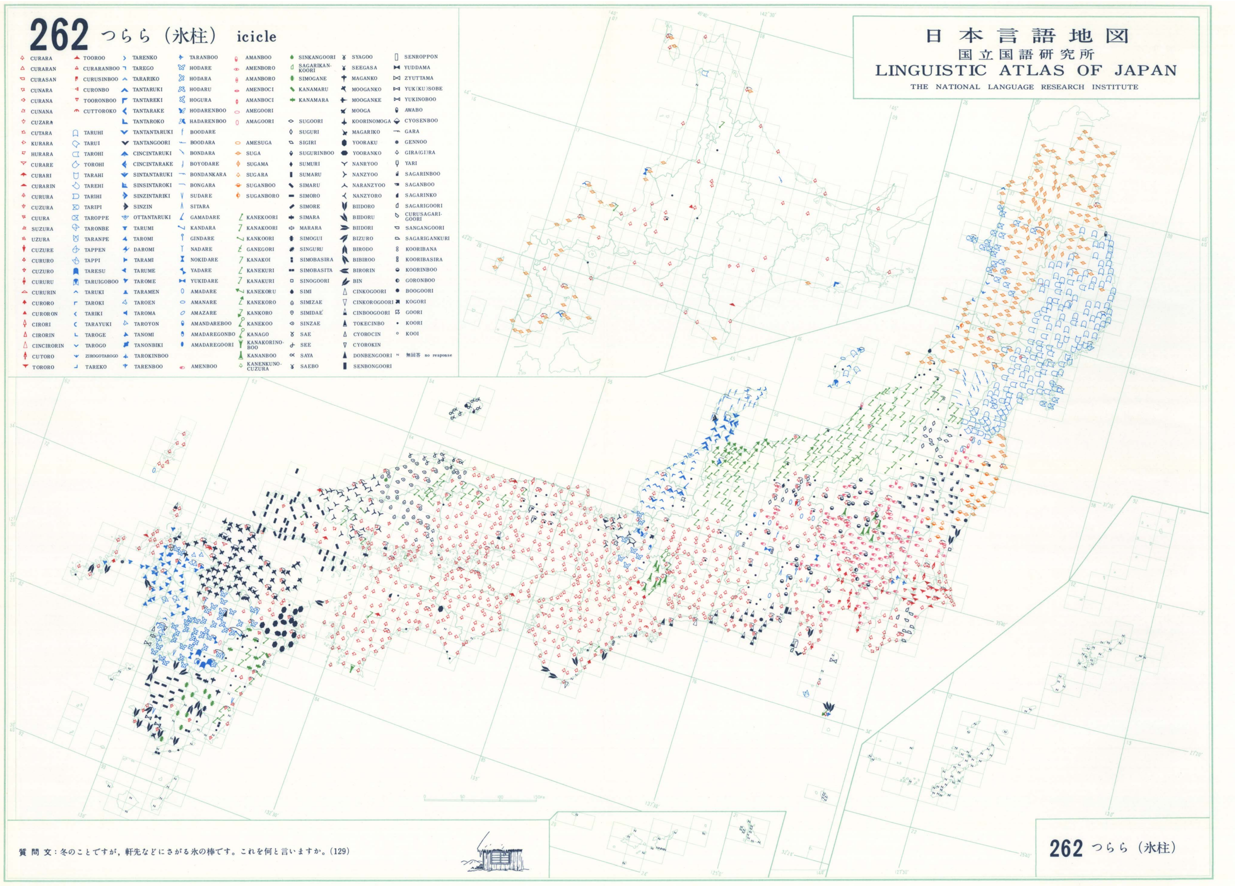


Figure 1: *LAJ Map #262: Icicle*

map of the LAJ depicts the highly disaggregated geographic distribution of pronunciations for (mostly) lexical items, accompanied by a legend that provides the complete list of all items, which are grouped based on their similarity. The underlying survey question that was used to elicit the responses depicted in each map is reported in the lower left corner of each map and accompanied by a graphical illustration (if applicable).[19]

[17] Onishi (2011) provides an detailed overview regarding the historical development of Japanese geolinguistics, and references the *Phonetic Dialect Atlas* with 29 maps from 1905 and the *Grammatical Dialect Atlas* (GDA) with 37 maps from 1906 as the oldest linguistic atlases on Japanese.

[18] See also Onishi (2011, pp.335-36) for a detailed discussion on how the distribution of pronunciations of the Japanese word for *icicle* in LAJ Map #262 can be interpreted as evidence in favor of the dialect radiation theory, according to which we would expect to find more modern forms (standard Japanese) in central regions and older variants in the country's periphery.

[19] Survey questions were intentionally kept as simple as possible to facilitate the understanding of respondents. The question underlying LAJ Map #262 reads: "冬のことですが、軒先などにさがる氷の棒です。これを何と言いますが？", which can be translated into: "*How do you call those ice sticks that in winter hang from eaves and such?*".

2.4 Other data

We use a Geographic Information System (GIS) to calculate the great-circle distance between market locations. To measure regional differences in climate conditions and soil quality, we utilize the Climate Mesh Data and the Land Classification Mesh Data provided by the National Land Planning Bureau of the Ministry of Land, Infrastructure, Transport and Tourism (MLIT).

3 Measuring the integration of fruit and vegetable markets

To assess the integration of wholesale markets for fruits and vegetables across Japan, we adopt the approach of Engel and Rogers (1996) and Crucini et al. (2010), who hypothesize that the price volatility of similar goods sold in different locations is related to a set of gravity variables (e.g. bilateral distance) separating these markets. Measuring the deviation from the law of one price as the log of the relative price of commodity c in market i to the price of the same commodity in market j at time t

$$\mathcal{P}_{ijc,t} \equiv \ln\left(P_{ic,t}\right) - \ln\left(P_{jc,t}\right) \tag{1}$$

allows us to compute the time series volatility of the LOP deviation as the standard deviation

$$REL\_PRC\_VOL_{ijc} \equiv \sqrt{\mathrm{Var}[\Delta\mathcal{P}_{ijc,t}]} \tag{2}$$

of the differenced series $\Delta\mathcal{P}_{ijc,t} = \mathcal{P}_{ijc,t} - \mathcal{P}_{ijc,t-1}$.

In the following we use the relative price volatility in Eq. (2) as our preferred measure for market integration, which according to Head and Mayer (2021, p. 39) is a "natural complement" to measuring economic integration based on bilateral trade volumes in the context of a gravity model. What renders the LOP approach of Engel and Rogers (1996) particularly useful in comparison to a quantity-based gravity approach (cf. Head and Mayer, 2014) is the possibility to also assess the integration of markets, which are primarily linked through the possibility of indirect arbitrage via a shared supplier network and therefore see little direct arbitrage in the form of market-to-market transactions.

In order to demonstrate that Japan's wholesale markets for fruits and vegetables are primarily linked through a shared supplier network, we rely on our highly disaggregated transaction-level data to compute the probability that any pair of wholesale markers are linked through direct versus indirect arbitrage.[20] Being able to link each market i's transactions in our data to an origin, which is either another wholesale market j or a supplier s located in one of Japan's 47 prefectures, we

[20] Japan's domestic market for fruits and vegetables is characterized by a strong comparative advantage, driven by the limited availability of agricultural land and relatively high transportation costs, which are further exacerbated by the perishable nature of most products (cf. Kano et al., 2013).

introduce the binary transaction indicators $T_{ijc,t} \in \{0,1\}$ and $T_{isc,t} \in \{0,1\}$. The commodity-specific market-to-market indicator $T_{ijc,t} \in \{0,1\}$ assumes a value of one whenever we observe a transaction value $X_{ijc,t} > 0$ between market j and market i in month t and zero otherwise. The analogously defined supplier-to-market indicator $T_{isc,t}$ assumes a value of one if we observed market i sourcing the value $X_{isc,t} > 0$ from supplier s in month t and zero otherwise. Referring to the sets of markets, suppliers and commodities by $\mathbb{J}$, $\mathbb{S}$ and $\mathbb{C}$, respectively, allows us to compute the probability

$$\mathrm{E}_i(T_{ijc,t}) = \frac{1}{120 \cdot |\mathbb{C}| \cdot (|\mathbb{J}| - 1)} \sum_t \sum_c \sum_{j \neq i} T_{ijc,t}, \tag{3}$$

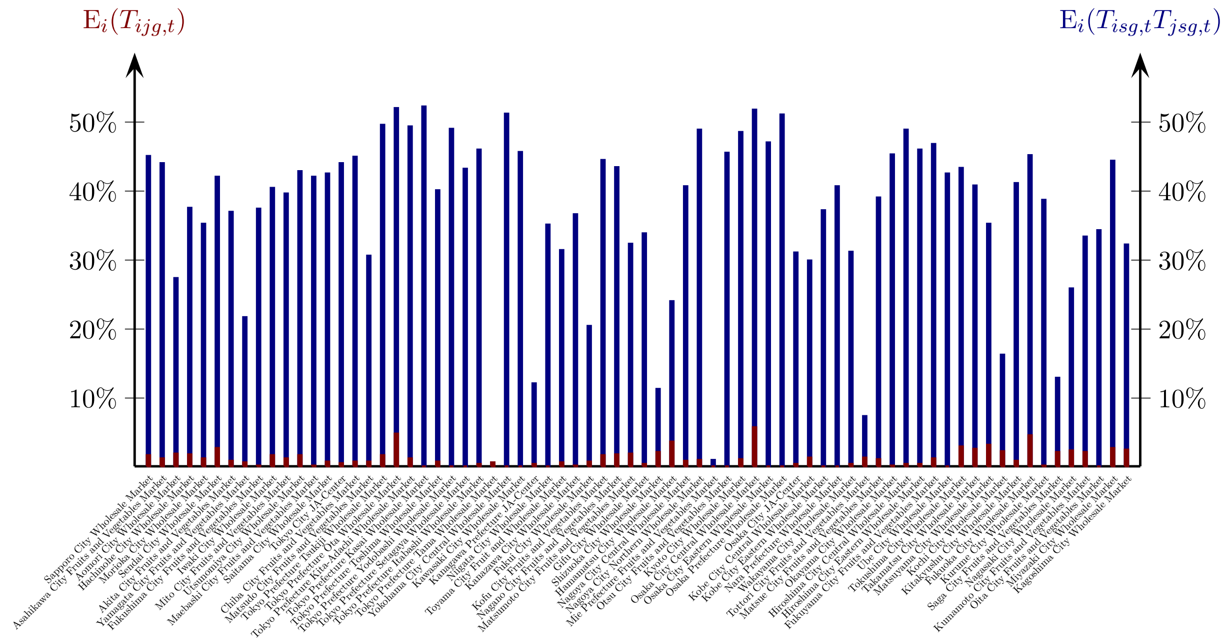
that market i is directly linked to other markets through bilateral arbitrage and the probability

$$\mathrm{E}_i(T_{isc,t}T_{jsc,t}) = \frac{1}{120 \cdot || \cdot |\mathbb{S}| \cdot (|\mathbb{J}| - 1)} \sum_t \sum_c \sum_s \sum_{j \neq i} T_{isc,t}T_{jsc,t} \tag{4}$$

that market i is indirectly linked to other markets through sharing a common supplier by averaging across all $t = 1, \ldots, 120$ months of our sample period.

Figure 2 compares the probabilities from Eqs. (3) and (4), using red bars to depict the probability $\mathrm{E}_i(T_{ijc,t})$ that markets are linked through direct arbitrage and blue bars to depict the probability $\mathrm{E}_i(T_{isc,t}T_{jsc,t})$ that markets are linked through the possibility of indirect arbitrage via a common supplier. Across all markets the chances of observing a direct market-to-market trans-

Figure 2: *How Japan's wholesale markets for fruits and vegetables are integrated*



Note: *Fig. 2 plots the average probability that a given market is linked to another market through a market-to-market transaction (in red) versus the average probability that a given market is linked to anther market through sharing a common supplier (in blue) across 72 Japanese wholesale markets for fruits and vegetables.*

action are close to zero. At the same time, we find that most markets in our sample are indirectly linked by sharing a common supplier.

To account for the role of shared supplier networks in explaining the volatility of LOP deviations from Eq. (2), we propose a new commodity-specific measure for shared supplier access

$$JOINT\_SUPPLIERS_{ijc} \equiv \frac{1}{120} \sum_{t} \sum_{s \in \mathbb{S}_{c,t}} w_{ijsc,t} T_{isc,t} T_{jsc,t}, \tag{5}$$

which aggregates the incidence $T_{isc,t}T_{jsc,t}$ that two markets i and j for commodity c are linked because they source in the same month t from a common supplier $s \in \mathbb{S}_{c,t}$ across all suppliers of commodity c at time t in set $\mathbb{S}_{c,t}$ and across all $t = 1, \ldots, 120$ months of our observation period, using the time-variant aggregation weights

$$w_{ijsc,t} \equiv \frac{X_{isc,t} + X_{jsc,t}}{\sum_{z \in \mathbb{S}_{c,t}} (X_{izc,t} + X_{jzc,t})} \tag{6}$$

to account for differences in the intensive margin of supplier-to-market transactions.

Finally, to illustrate the usefulness of the relative price volatility as a measure for market integration, that delivers results comparable to those of gravity estimations based on bilateral trade flows (cf. Head and Mayer, 2014), we extend the empirical specification of Engel and Rogers (1996) to the commodity level, and estimate

$$\begin{aligned} REL\_PRICE\_VOL_{ijc} = \alpha &+ \beta \ln DIST_{ij} + \gamma\, JOINT\_SUPPLIERS_{ijc} \\ &+ \delta_{PREF} PREF\_BORD_{ij} + \delta_{REG} REG\_BORD_{ij} + \delta_{ISL} ISL\_BORD_{ij} \\ &+ \sum_{n=1}^{72} \mu_n D_n + \sum_{c=1}^{31} \nu_c D_c + \varepsilon_{ijc}. \end{aligned} \tag{7}$$

In Eq. (7) the volatility of the market-pair specific relative price for commodity c is regressed on the logarithm of geographic distance $\ln DIST_{ij}$ between market i and market j, our commodity-specific measure of joint supplier access $JOINT\_SUPPLIERS_{ijc}$, a set of prefecture-, region- and island-specific border dummies $PREF\_BORD_{ij} \in \{0,1\}$, $REG\_BORD_{ij} \in \{0,1\}$ and $ISL\_BORD_{ij} \in \{0,1\}$, which take a value of one if the location of markets i and j differs at the level of 46 prefectures, 8 regions and 4 main islands, as well as two sets of market- and commodity-specific fixed effects $D_n \in \{0,1\}$ and $D_c \in \{0,1\}$, and a random error term ε_{ijc}.[21] Following the argumentation of Engel and Rogers (1996), we impose market-specific fixed effects to account for level differences in the volatility of local prices, that may result either from idiosyncratic measurement error or from

[21] See Table A.2 (delegated to the Appendix), for a list of Japan's 46 prefectures (excluding Okinawa), which can be aggregated into 8 regions (Hokkaidô, Tôhoku, Kantô, Chûbu, Kansai, Chûgoku, Shikoku, and Kyûshû) and 4 islands (Hokkaidô, Honshû, Shikoku, Kyûshû).

seasonalities. Commodity-specific fixed effects are included to absorb volatility differences across commodities that differ among others in terms of their seasonality, storability and transportation costs.

Table A.4 (presented in the Appendix) reports the results from estimating Eq. (7), confirming the anticipated positive and highly significant relationship between geographic distance and the volatility of relative market prices, consistent with Engel and Rogers (1996). Importantly, and central to our analysis, we find that markets connected by one or more common suppliers exhibit a significantly lower relative price volatility. Comparing Columns (1) and (2) of Table A.4 moreover reveals that accounting for joint suppliers reduces the effect of geographic distance by approximately 40%. This result suggests that Japan's wholesale markets for fruits and vegetables are integrated not primarily through direct arbitrage but through indirect arbitrage facilitated by a network of shared suppliers.

4 Estimating localized tastes

In order to recover localized tastes in the vicinity of 72 Japanese wholesale markets for fruits and vegetables, we associate each market with the nearest municipality (see Table A.2) covered by the Japanese Family Income and Expenditure Survey. Doing so leaves us with 61 municipalities (index by mnemonic m) for which we can use detailed micro-level household data from the Japanese Family Income and Expenditure Survey to estimate municipality-specific taste parameters for each of the 31 commodities covered by our analysis.

Taste estimation follows Atkin (2013) and Colson-Sihra et al. (2023), who rely on the Almost Ideal Demand System (AIDS) (cf. Deaton and Muellbauer, 1980) to identify commodity-specific localized tastes as residual demand shifters taking into account the effects of income, prices, and household characteristics.[22] The AIDS's functional form flexibility provides several key advantages but comes with the drawback of managing the complexity associated with high product dimensionality.[23] Because the number of cross-price elasticities that need to be estimated increases squarely in the number of commodities, we follow Nevo (2011) and specify a multi-level demand system, which distinguishes between a upper level with a broad commodity classification (indexed by mnemonic b) and a lower level with a commodity-specific index c (see Section 2.2 for

[22]See also Fajgelbaum and Khandelwal (2016), Hummels and Lee (2018), Hillrichs and Vannoorenberghe (2022) Feenstra and Hong (2022) and Anderson and Zhang (2025) for recent applications of the Almost Ideal Demand System outside its original industrial organization context.

[23]Colson-Sihra et al. (2023) highlight three key advantages that render the Almost Ideal Demand system particular useful for the estimation of localized tastes: (*i.*) demand functions derived from the AIDS are first-order approximations to any set of demand functions derived from utility maximization; (*ii.*) the AIDS accomodates heterogeneous cross-price elasticities and nonhomothetic consumption pasterns; and (*iii.*) it allows for localized tastes that are additively separable from price and income effects.

the definition of the broad commodity classification).

As in Atkin (2013) and Colson-Sihra et al. (2023) the AIDS minimum expenditure function for household h

$$\ln e(\boldsymbol{\tau}, \mathbf{p}_{hg}, u) = \text{Const.} + \sum_k \tau_{mg} \ln p_{hg} + \frac{1}{2} \sum_g \sum_{g'} \theta_{gg'} \ln p_{hg} \ln p_{hg'} + u\phi_0 \prod_g p_{hg}^{\phi_g} \;\; \forall \;\; g \in \{b, c\}, \quad (8)$$

allows the first-order price terms to vary by location (in our case by municipality m). The expenditure function in Eq. (8) defines the minimum expenditure $e(\boldsymbol{\tau}_{mg}, \mathbf{p}_{hg}, u)$ that a household h with taste vector $\boldsymbol{\tau}_{mg}$ has to spend in order to reach the desired utility level u at a given price vector $\mathbf{p}_{hg}$ for good $g \in \{b, c\}$. The parameters τ_{mg}, $\theta_{gg'}$, and ϕ_g satisfy (*i.*) the adding-up constraint ($\sum_g \tau_{mg} = 1$), (*ii.*) homogeneity ($\sum_g \theta_{gg'} = \sum \phi_g = 0$), and (*iii.*) symmetry ($\theta_{gg'} = \theta_{g'g} \;\forall\; g, g'$).

As outlined in Nevo (2011), Shephard's lemma can now be applied to Eq. (8) in order to derive the commodity-level demand function in budget shares

$$s_{hc} = \tau_{mc} + \sum_{c'} \theta_{cc'} \ln p_{hc'} + \phi_c \ln \left(\frac{X_{hb}}{P_{hb}} \right), \quad (9)$$

in which s_{hc} is household h's budget share for commodity c in the broad category b, τ_{mc} is the municipality-specific taste parameter we seek to estimate, $\ln p_{hc'}$ is the log price that household h paid for commodity c, and X_{hb}/P_{hb} is the household's real expenditure X_{hb}/P_{hb} within the broad category b with P_{hb} as the corresponding AIDS price index.

Analogously, the demand function in budget shares for broad commodity categories (i.e. $g = b$) can be derived as

$$s_{hb} = \tau_{mb} + \sum_{b'} \theta_{bb'} \ln P_{hb'} + \phi_b \ln \left(\frac{X_h}{P_h} \right) + \boldsymbol{\zeta} \mathbf{H}'_h, \quad (10)$$

in which τ_{mb} denotes the municipality-specific taste parameter for the broad commodity category b, X_h/P_h is household h's real income with P_h as the AIDS price index computed from the commodity specific prices $\ln p_{hc}$ in Eq. (9), and $\mathbf{H}_h$ is a vector of variables that shift demand.

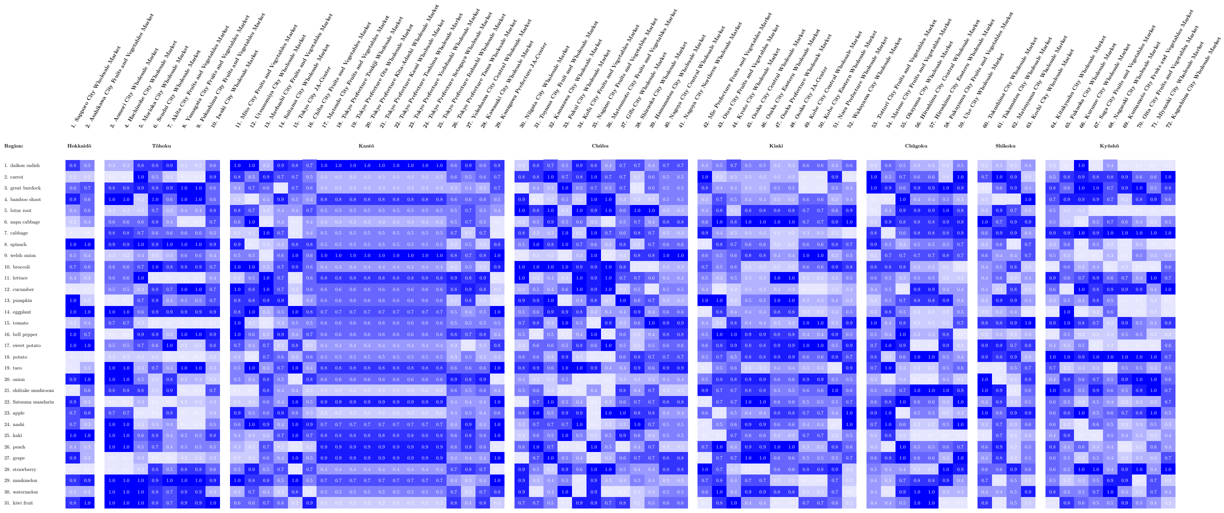
Since, we are primarily interested in obtaining municipality-specific taste estimates at the commodity level $\hat{\tau}_{mc}$, we estimate household h's demand for commodity c in municipality based on Eq. (9), which is transformed into the following estimation equation:

$$s_{hct} = \tau_{mc} + \sum_{c'} \theta_{cc'} \ln p_{mc't} + \phi_c \ln \left(\frac{X_{hbt}}{P^{\star}_{mbt}} \right) + \boldsymbol{\xi} \mathbf{H}'_{ht} + z_t + \varepsilon_{hct}. \quad (11)$$

Instead of using the commodity price $p_{hc't}$ that household h reports in the Japanese Family Income and Expenditure Survey, the median price of commodity c across all wholesale market in municipality m is used for estimation (cf. Atkin, 2013; Atkin et al., 2021; Colson-Sihra et al., 2023).[24] To allow for linear estimation of Eq. (11), the AIDS price index P_{hc} from Eq. (9) is replaced by the Stone price index $P^{\star}_{mbt} \equiv \sum_{c'} s_{mc't} \ln p_{mc't}$ for the broad product category b in municipality m, in which $s_{mc't}$ is the share of municipality-wide expenditure on good c' in the broad product category b (cf. Nevo, 2011; Deaton and Muellbauer, 1980). Following Subramanian and Deaton (1996), a vector of household characteristics $\mathbf{H}_{ht}$ (demographics as well as occupation- and industry-specific fixed effects) is introduced to account for unobserved heterogeneity that may simultaneously influence the demand and the taste for a specific commodity.[25] A set of fixed effects z_t controlling for the calender month is introduced to account for seasonality effects. The error term is denoted by ε_{hct}.

We report the OLS results from estimating Eq. (11) in the Tables A.5 to A.8 (delegated to the Appendix) and focus here on visualizing the geographic distribution of the estimated taste parameters, $\hat{\tau}_{mc}$. In Table 1 the percentiles of the taste estimates for 31 fruits and vegetables are

Table 1: *Distribution of commodity-specific taste estimates across wholesale markets*



Notes: Table 1 plots the percentiles of the taste estimates for 31 fruits and vegetables across 72 wholesale markets, which are ordered from North-West (1. Sapporo City Wholesale Market) to South-East (72. Kagoshima City Wholesale Market) and grouped into Japan's eight administrative regions (Hokkaidô, Tôhoku, Kantô, Chûbu, Kinki, Chûgoku, Shikoku and Kyûshû). Darker (lighter) tones of blue are associated with higher (lower) percentiles.

reported across 72 wholesale markets, which are located in 61 distinct municipalities.[26] Darker shades of blue in Table 1 represent higher percentiles, while lighter shades indicate lower percentiles. Given Japan's elongated, banana-shaped geography – spanning over 1,800 km from the

[24] Household prices in the Japanese Family Income and Expenditure Survey are only erratically reported and most likely measured with error. Colson-Sihra et al. (2023) replace household prices by city-level median prices because these prices are robust to outlines and not confounded by the household's quality choice.

[25] We account for the total number of household members, the number of members under 18, those over 65, the age and employment status of the household head, and two distinct sets of fixed effects specific to occupation and industry.

[26] Because tastes are estimated at the municipality level, wholesale markets that are located in the same municipality share the same tastes. Our sample features four municipalities (Nagoya, Osaka, Kobe and Hiroshima) with two wholesale markets and one municipality (Tokyo) with eight wholesale markets (see also Table A.2).

northeast to the southwest – the wholesale markets in Table 1 are ordered geographically, starting from the northwest (1. Sapporo City Wholesale Market) to the southeast (72. Kagoshima City Wholesale Market). Consequently, the column numbering in Table 1 provides a rough indication of the geographic distances between the wholesale markets. Additionally, this ordering allows for clustering wholesale markets by their location within one of Japan's eight administrative regions: Hokkaidô, Tôhoku, Kantô, Chûbu, Kinki, Chûgoku, Shikoku and Kyûshû.

The following three key insights follow from Table 1: (*i.*) Comparing the columns in Table 1 shows that estimated tastes are highly localized, with markets in the same administrative region or close proximity often displaying similar taste profiles. (*ii.*) While some commodities, such as daikon radish (first row of Table 1), display distinct northwest-versus-southeast taste patterns, there is no consistent monotonic relationship between taste similarity and geographic distance. (*iii.*) Examining the rows of Table 1 reveals significant heterogeneity in the spatial distribution of commodity-specific tastes. This variation provides a valuable source of data for analyzing how taste heterogeneity influences the volatility of LOP deviations at the market-pair-by-commodity level (see Section 5).

Following Colson-Sihra et al. (2023), the commodity-specific taste estimates $\hat{\tau}_{mc}$, that results from estimating Eq. (11), now can be used to construct the market-pair-specific taste distance

$$TASTE\_DIST_{ijc} \equiv \left| \hat{\tau}_{ic} - \hat{\tau}_{jc} \right| , \tag{12}$$

which allows us to capture commodity-specific preference differences between the markets i and j.

5 Testing the localized tastes hypothesis

5.1 Identification strategy

According to the localized tastes hypothesis, a lack of bilateral market integration – measured by high volatility in LOP deviations – stems from historically determined regional taste differences that persist even after their original causes have faded. To test this hypothesis, we propose a straightforward two-stage estimation procedure: At the first stage, we use the regional overlap in historical dialects (cf. Falck et al., 2012) as a proxy for past cultural exchange, to isolate the persistent variation in bilateral taste dissimilarity. At the second stage, we then use these predicted regional taste differences to identify the causal effect of bilateral taste heterogeneity on the integration of regional wholesale markets, as reflected in the volatility of LOP deviations.

To obtain our second-stage specification, we extend Eq. (7) to account for the effect of commodity-specific regional taste differences $TASTE\_DIST_{ijc}$ on the volatility of the market-pair specific relative price for commodity c

$$\begin{aligned} REL\_PRICE\_VOL_{ijc} = \alpha + \beta\, TASTE\_DIST_{ijc} + \gamma\, JOINT\_SUPPLIERS_{ijc} \\ + \mathbf{X}'_{\mathbf{ij}}\boldsymbol{\delta} + \sum_{n=1}^{72} \mu_n D_n + \sum_{c=1}^{31} \nu_c D_c + \varepsilon_{ijc}, \end{aligned} \tag{13}$$

with vector $\mathbf{X_{ij}}$ summarizing all market-pair-specific controls from Eq. (7). In order to instrument the estimated taste distance $TASTE\_DIST_{ijc}$ from Eq. (12) we estimate the first-stage regression

$$TASTE\_DIST_{ijc} = \rho + \sigma\, HIST\_DIALECT\_SIM_{ij} + \mathbf{X}'_{\mathbf{ij}}\boldsymbol{\theta} + \sum_{n=1}^{72} \xi_n D_n + \sum_{c=1}^{31} \zeta_c D_c + \epsilon_{ijc}, \tag{14}$$

in which the estimated taste distance $TASTE\_DIST_{ijc}$ from Eq. (12) is regressed on a historical dialect similarity index $HIST\_DIALECT\_SIM_{ij}$ (which is explained in more detail below) as well as on the vector $\mathbf{X_{ij}}$ of geographic controls (cf. Jeszenszky et al., 2019) and the complete sets of market- and commodity-specific fixed effects $D_n \in \{0, 1\}$ and $D_c \in \{0, 1\}$ followed by the random error term ϵ_{ijc}.

Our identification strategy builds on three key insights from the tastes literature, which studies the emergence, transmission and persistence of local tastes (cf. Maystre et al., 2014; Bisin and Verdier, 2014; Atkin, 2013, 2016; Colson-Sihra et al., 2023). Summarizing earlier contributions to this literature, Head and Mayer (2013) conclude that:

(*i.*) Localized food preferences originally developed in response to the abundant supply of local varieties, before the advent of cheap, refrigerated transportation;

(*ii.*) Cultural exchange facilitated by economic interactions (e.g. trade and migration) lead to a diffusion/convergence of historically shaped regional tastes;

(*iii.*) Localized tastes persist even after their original determinants (e.g. prohibitive transportation costs) have diminished in importance.

Atkin (2013) provides systematic evidence on point (*i.*) by showing that the formation of food tastes in India is driven by the heterogeneous agro-climatic endowments of Indian regions.[27] For Japan, whose traditional dietary cultures (*washoku*) have been inscribed as Intangible Cultural Heritage of Humanity in 2013 (cf. UNESCO, 2013), anecdotal evidence on the origins of localized tastes is collected by the Ministry of Agriculture, Forestry and Fisheries (MAFF), which has put together a comprehensive list of Traditional Foods in Japan (cf. MAFF, 2025b).[28]

[27] Colson-Sihra et al. (2023) focus on France, and document a distinct northwest/southeast divide in the use of butter versus olive oil, which can be traced back to a historical map from the 1950s, and a geographic divide between wine consumption (in southern France), cidre consumption (in northwestern France) and beer consumption (in northern France), which can be traced back to a historical map from the 1850s.

[28] To qualify as a traditional food, a product must have origins in a specific region, either through locally developed

According to Head and Mayer (2013), trade and migration are the only factors capable of counteracting the natural divergence of local tastes. Increased trade openness exposes consumers to new and often cheaper product varieties, while migration introduces preferences from other regions, making individuals more likely to overcome their initial resistance to unfamiliar tastes.[29] Supportive evidence on point (*ii.*) is provided by Kónya and Ohashi (2007), Aizenman and Brooks (2008) and Bargain (2024), who link the homogenization of consumption patterns to the integration of goods markets, as well as by Bronnenberg et al. (2012), Atkin (2016) and Hut (2020), who demonstrates that migrants tend to retain their preferences when settling in new locations.[30]

Evidence on point (*iii.*) comes from Oster (2018) and Hut and Oster (2022), who show that food consumption patterns are highly persistent and insensitive to change. Staehle (1934), Bronnenberg et al. (2012), Atkin (2016) and Hut (2020) provide complementary evidence on the persistence of food preferences of migrants.

Motivated by these three stylized facts, we propose an instrumental variable strategy, that uses bilateral variation in historical cultural similarity to predict persistent differences in contemporaneous tastes. This approach is based on the notion that regions with more intensive exchange in the past gradually converged in their food cultures and, ultimately, in their tastes. Because historical taste (dis)similarities that emerged from evolutionary processes (cf. Bisin and Verdier, 2001; Atkin, 2013; Maystre et al., 2014) in response to past cultural exchange are expected to persist over time, we can rely on this historical variation to account for a potential measurement error the contemporaneous taste distance and to purge it from the influence of possible contemporaneous confounders.

To measure the intensity of past cultural exchange between locations in Japan, we adopt the approach of Falck et al. (2012) and Lameli et al. (2015), using historical dialect data from the Linguistic Atlas of Japan Database (introduced in Subsection 2.3) to construct a historical dialect similarity index.[31] As pointed out by Falck et al. (2012, pp. 228-31), historical dialects are the outcome of a long-term evolutionary process (cf. Schmidt, 2010) that reflects the entirety of past

recipes or the use of unique ingredients, with a history traceable to before World War II (cf. MAFF, 2025b). Japanese pickles (*tsukemono*) are an formidable example for how regional heterogeneity (e.g. in terms of climate and remoteness from the coast) contributed to the emergence of local food cultures (cf. MAFF, 2025a). Kimura (2021) reports that at least 46% of the 35 vegetable varieties recognized by Kyôtô Prefecture as "Kyôtô traditional vegetables" are predominantly or exclusively used for producing pickled vegetables (*tsukemono*).

[29] Maystre et al. (2014) propose a model in which trade integration leads to cultural convergence because increased trade exposure changes the incentives of parents to pass on their preferences to the next generation. A systematic review of the theoretical literature on the relationship between globalization and cultural diversity is provided by Bisin and Verdier (2014).

[30] See Ochsner and Roesel (2020) for evidence on migrants that carry their political preferences with them.

[31] In order to capture the effect of cultural proximity on international trade at the country-pair level researchers have focused on ethnic networks (cf. Rauch and Trindade, 2002), bilateral opinions (cf. Disdier and Mayer, 2007), bilateral trust (cf. Guiso et al., 2009), colonial linkages (cf. Head et al., 2010), voting in the Eurovision Song Context (cf. Felbermayr and Toubal, 2010), a multidimensional cultural proximity index (cf. Möhlmann et al., 2010), linguistic proximity (cf. Melitz and Toubal, 2014), virtual proximity, (cf. Hellmanzik and Schmitz, 2015) and genetic distance (cf. Bove and Gokmen, 2018).

interactions, and therefore are interrelated with historical religious and political borders, ancient transportation networks and past migration flows. We argue that these and other historical interactions, which left some long-lasting imprints that are encapsulated in local dialects, not only shaped the distribution of historical dialects but also influenced evolutionary convergence/divergence processes (cf. Hinskens et al., 2005) that were responsible for the formation of regional taste differences that persist until today. Using historical dialect similarity as a proxy for past cultural exchange in the first-stage regression from Eq. (14) therefore allows us to predict the persistent component of the contemporaneous taste distance from Eq. (12).

In order to compute our historical dialect similarity index, we match the individual responses to 141 questionnaire items from the Linguistic Atlas of Japan Database (see Table A.3) to the exact coordinates of up to 2,400 survey locations. The following example based on LAJ survey item #129 *icicle* [氷柱] (see LAJ Map #262 in Figure 1) illustrates our approach: The LAJ lists at total of 232 historical pronunciations for the Japanese word for *icicle* (e.g. "CURARA", "TARUHI", "BOODARE", "KANEKOORI", ...). By matching the distribution of historical dialects from Figure 1 to 25-km-radius circles around the locations of wholesale markets, we can associate each market location i with a set $\mathbb{A}_{qi}$ of distinct answers to questionnaire item $q \in \mathbb{Q}$. Repeating this process for all 141 questionnaire items of the LAJDB included in the set $\mathbb{Q}$ allows us to characterize the historical dialect at location i through the sets of distinct answers $\mathbb{A}_{qi}\ \forall\ q \in \mathbb{Q}$.

To proxy the intensity of past cultural exchange between two markets, we compute the overlap in market-specific historical dialects using the Jaccard index (cf. Jaccard, 1901), which is an established similarity measure in the field of linguistic geography.[32] For a given question q and market-specific sets of distinct answers $\mathbb{A}_{qi}$ and $\mathbb{A}_{qj}$ the similarity in dialects between a given pair of wholesale markets i and j can be measured by

$$J_{qij} = J_{qji} = \frac{|\mathbb{A}_{qi} \cap \mathbb{A}_{qj}|}{|\mathbb{A}_{qi} \cup \mathbb{A}_{qj}|} \in [0,1], \tag{15}$$

in which $|\mathbb{X}|$ denotes the the cardinality of set $\mathbb{X}$. Averaging across all questions $q \in \mathbb{Q}$ of the LAJDB then yields the average overlap in market-specific historical dialects

$$HIST\_DIALECT\_SIM_{ij} = \frac{1}{|\mathbb{Q}|} \sum_{q \in \mathbb{Q}} J_{qij} \in [0,1], \tag{16}$$

which we use as a proxy for past cultural exchange between any pair of market locations i and j.

To visualize the similarity of historical dialects and demonstrate how their geography reflects

[32] In order to measure the overlap in the responses to the 66 survey items of the linguistic atlas of the German Empire Falck et al. (2012) compute a binary measure that takes a value of one if the most frequent pronunciation in two regions is the same and zero otherwise.

past interactions, we plot in Table A.9 (delegated to the Appendix) the historical dialect similarity matrix for all $i \times j$ market pairs. As in Falck et al. (2012, pp. 229-31), it is then possible to trace the existence of language enclaves back to past events of mass migration. Examining the dialects of Japan's two northernmost wholesale markets – Sapporo and Asahikawa, both on the island of Hokkaidô – we find little overlap with dialects from northeastern Japan but a striking resemblance to those spoken in the central Kantô region. The historical explanation for this pattern in Table A.9 is Hokkaidô's systematic colonization (cf. Mason, 2012), which was driven by large-scale internal migration primarily from central Japan (cf. Shigeaki, 1991). Complementary evidence on the role of geography is provided by Jeszenszky et al. (2019), who explore the effects geographic distance and administrative boundaries on the similarity of Japanese dialects, and Onishi (2011), who shows that the spatial distribution of negative verbal suffixes (*nai* versus *-n*) correlates with Japan's surface topography.

Our primary concern regarding the estimated effect of taste heterogeneity on the volatility of LOP deviations in Eq. (13) is the potential for attenuation bias due to measurement error (cf. Fuller, 1987; Hausman, 2001).[33] While regional taste differences cannot be directly observed, they can be estimated – although potentially with some margin of error. By combining the estimated taste coefficients from Section 4 according to the assumed functional form from Eq. (12), we hence derive a bilateral taste distance measure that may be subject to measurement error, necessitating the use of an instrumental variable approach (cf. Aw et al., 2023).

While we cannot entirely rule out omitted variable bias and reverse causality as potential sources of endogeneity, we are less concerned about unobserved heterogeneity. This is because market- and commodity-specific trends are accounted for through appropriately specified fixed effects in Eq. (13). Moreover, our results remain robust even when incorporating a more comprehensive fixed effects structure that captures bilateral variation (see Subsection 6.1).

Reverse causality is a concern if improved market integration – reflected by lower volatility in LOP deviations – leads to the homogenization of regional tastes. Colson-Sihra et al. (2023) examine changes in regional food preferences in France from 1974 to 2005, based on household survey data. They find that geographically closer regions shared more similar tastes in 1974 – a geographic pattern that had disappeared by 2005. However, their analysis finds no evidence that improved bilateral market integration drove this decline in the influence of geographic distance on taste differences. Instead, the evolution of regional tastes in France seems to be shaped more by sociocultural factors, such as educational composition, than by geography.

In summary, our instrumental variable approach is expected to be most effective in addressing a

[33] For meta-level evidence on the pervasiveness and quantitative importance of measurement error see Jiang (2017), Lal et al. (2024) and Pancost and Schaller (2024).

potential attenuation bias. According to the localized tastes hypothesis, regional taste differences should have a positive effect on the volatility of relative prices. However, in the presence of measurement error, this effect would be systematically underestimated by OLS, potentially biasing the results toward zero.

The validity of our instrumental variable approach hinges on the plausibility of the exclusion restriction, which requires that the volatility of LOP deviations depends on our historical dialect similarity index solely through its effect on contemporaneous taste differences. This restriction would be violated if markets sharing a common historical dialect experience higher bilateral trust (cf. Guiso et al., 2009) or are connected through bilateral migration and trade networks (cf. Falck et al., 2012; Bauernschuster et al., 2014; Lameli et al., 2015) that influence the volatility of relative market prices. However, because Japanese wholesale markets are not directly integrated through bilateral arbitrage but rather through indirect linkages via a network of shared suppliers (see Figure 2 and Table A.4), we argue – and provide auxiliary evidence – that such alternative channels, which may operate at the market-pair level, do not violate our exclusion restriction (see Subsection 6.2).

Subsection 5.1 introduces a novel instrumental variable strategy which relies on variation in historical dialects to identify the effect of taste heterogeneity on the volatility of LOP deviations. Inspired by Head and Mayer's (2013) localized tastes hypothesis, the instrument uses a well-established proxy for past cultural exchange to predict historically determined regional taste differences that remain relevant today. We argue that this instrument effectively addresses attenuation bias caused by measurement error in contemporaneous taste estimates and is plausibly exogenous, given the structure of Japanese wholesale market integration.

5.2 Results

Table 2 presents the results from estimating Eq. (13) using OLS and 2SLS. In line with the localized tastes hypothesis, Column (1) shows that a greater taste distance is associated with higher volatility in LOP deviations. However, the OLS estimate in Column (1) is small and only marginally significant. Gradually adding the controls from Section 3 in the columns (2) to (4), further diminishes the magnitude of the taste heterogeneity effect, which ultimately becomes statistically indistinguishable from zero.

Our main concern regarding a possible endogeneity of the OLS estimates from Columns (1) to (4) relates to attenuation bias caused by measurement error. The taste distance in Eq. (12) is constructed from taste parameters that are unobservable and therefore have to be estimated (see Section 4). Measurement error that results from the estimation of the underlying taste parameters therefore will bias the OLS estimates in Table 2 towards zero, which explains why the taste estimate in our preferred OLS specification from column (4) is relatively small and statistically insignificant.

To overcome a possible attenuation bias and to address other endogeneity concerns, taste distance in Column (5) of Table 2 is instrument using historical dialect similarity (see Subsection 5.1). Reassuringly, the effect of taste heterogeneity in Column (5) is substantially larger than under OLS and statistically different from zero at a 5% level of significance. The 1st-stage results from Table 2 show that historical dialect similarity has a negative and highly significant effect on the market-pair-specific taste distance, supporting the idea that past cultural exchange contributed to a convergence of regional tastes (cf. Head and Mayer, 2013). A p-value of .0006 for the Kleibergen and Paap (2006) test statistic and a 1st-stage F-statistic of 18.42 confirm the relevance and strength of historical dialect similarity as an instrument for contemporaneous taste dissimilarity. Interestingly, the 1st-stage results from Table 2 suggest that the distance in regional tastes is uncorrelated with our commodity-specific measure of joint supplier access from Section 3, which captures the indirect integration of wholesale markets though sourcing from common suppliers. We interpret this finding as suggestive evidence against a possible reversed causality problem, which would arise when better integrated markets converge in terms of their local tastes.[34] To assess the quantitative importance of taste heterogeneity for the volatility of LOP deviations in Table 2, our preferred point estimate of 1.0401 is evaluated at the sample mean taste distance .02387, which adds 0.0248 to the volatility in relative prices. Comparing this change to the sample mean of .2878 for the volatility of LOP deviations, we conclude that eliminating regional taste heterogeneity would reduce the volatility in relative prices by approximately 9%. For comparison, setting the average geographic distance between markets to zero, would result in a 12% decline in the volatility of LOP deviations relative to the sample mean. Similarly, assuming the supplier network variable to reach its theoretical maximum of one, would decrease the relative price volatility by about 7% relative to the sample mean. These findings indicate that a substantial and economically meaningful portion of the volatility of LOP deviations can be attributed to differences in regional preferences.

[34] See also Colson-Sihra et al. (2023), who argue that concerns about cultural homogenization due to market integration are unfounded. By analyzing the distribution of regional tastes in France between 1974 and 2005, they conclude that the weakening influence of geographic distance on taste dissimilarity cannot be solely attributed to bilateral factors such as the ability to move goods, people and ideas. Instead, their findings suggest that regional taste differences are more closely linked to sociocultural similarities between regions (e.g. in terms of educational and nationality composition). In the Online Appendix (available upon request), we replicate the analysis by Colson-Sihra et al. (2023) using household expenditure data from Japan's Family Income and Expenditure Survey covering the years 1985 to 2019. For approximately half of the commodities in our sample, we find that the effect of geographic distance on taste dissimilarity weakens somewhat over time, while no consistent time trends can be identified for the remainder of our sample.

Table 2: *Testing the localized tastes hypothesis*

Dependent variable: Relative price volatility for commodity c b/w the markets i and j ($REL\_PRICE\_VOL_{ijc}$)					
2nd-stage results					
Model:	**OLS-FE**				**2SLS-FE**
Specification:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$TASTE\_DIST_{ijc}$	0.1575*	0.0810	0.0739	0.0741	1.0401**
	(.0797)	(.0660)	(.0651)	(.0650)	(.4122)
$\ln DIST_{ij}$		0.0146***	0.0088***	0.0094***	0.0071***
		(.0013)	(.0014)	(.0016)	(.0021)
$JOINT\_SUPPLIERS_{ijc}$			−0.0641***	−0.0631***	−0.0615***
			(.0132)	(.0135)	(.0136)
$PREF\_BORD_{ij}$				−0.0049**	−0.0104***
				(.0023)	(.0022)
$REGION\_BORD_{ij}$				−0.0023*	−0.0021*
				(.0012)	(.0012)
$ISLAND\_BORD_{ij}$				0.0053**	0.0037
				(.0025)	(.0028)
Fixed effects:					
Market-specific fixed effects:	✓	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓	✓
Summary statistics:					
# of observations:	158,472	158,472	158,472	158,472	158,472
R^2:	.8153	.8183	.8209	.8209	-
Kleibergen-Paap (p-value):	-	-	-	-	.0009
1st-stage results					
$HIST\_DIALECT\_SIM_{ij}$					−0.0149***
					(.0037)
$\ln DIST_{ij}$					0.0018***
					(.0005)
$JOINT\_SUPPLIERS_{ijc}$					−0.0018
					(.0018)
$PREF\_BORD_{ij}$					0.0038***
					(.0007)
$REGION\_BORD_{ij}$					−0.0006
					(.0006)
$ISLAND\_BORD_{ij}$					0.0019*
					(.0011)
Fixed effects:					
Market-specific fixed effects:					✓
Commodity-specific fixed effects:					✓
Summary statistics:					
# of observations:					158,472
R^2:					.2500
F-statistic (p-value):					.0003

Notes: Robust standard errors clustered at the commodity level; significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6 Robustness analysis

6.1 Accounting for unobserved heterogeneity

In Subsection 5 it has been argued that the difference between the small and insignificant OLS results and the sizable, significant 2SLS results from Table 2 is consistent with attenuation bias caused by measurement error in the bilateral taste distance. To rule out omitted variable bias as an alternative explanation for this result, Table 3 allows for a more complex fixed effect structure than in Table 2, incorporating island-pair-, region-pair- and prefecture-pair-specific fixed effects.[35] Because the 72 wholesale markets for fruits and vegetables covered by our analysis are located across the 4 major islands (Hokkaidô, Honshû, Shikoku and Kyûshû), which are further divided into 8 administrative regions and a total of 46 prefectures (see Table A.2), it is possible to absorb all island-, region- and prefecture-pair-specific unobserved heterogeneity through a set of appropriately specified fixed effects.[36] As a result, the effects of all market-integrating forces, which plausibly can be expected to vary at a more aggregate level than between individual cities (e.g. transportation infrastructure or networks), are largely absorbed, which substantially reduces the potential for an endogeneity problem due to omitted variables.[37]

Column (1) of Table 3 replicates our baseline result from Table 2. Introducing island-pair- and region-pair-specific fixed effects in Columns (2) and (3) does not affect the magnitude or significance of the taste distance effect. Reassuringly, the 1st-stage results in Columns (2) and (3) also resemble those from our baseline specification in Column (1), confirming the relevance and strength of our instrument. It is worth noting, that the geographic distance coefficient decreases by about two-thirds and becomes statistically insignificant, when region-pair-specific fixed effects are included in Column (3). A plausible explanation for this finding is that the region-pair-specific variation of an omitted variable – previously captured by the distance coefficient in Columns (1) and (2) – is now fully absorbed by region-pair-specific fixed effects. Adding a large number of prefecture-pair-specific fixed effects in Column (4) reduces the power of our instrument, which is reflected by a considerably smaller 1st-stage F-statistic of 7.40. According to Stock and Yogo (2005), historical dialect similarity in Column (4) therefore qualifies as a weak instrument, necessitating weak-instrument robust inference. Both, the Anderson and Rubin (1949) and the Stock and Wright (2000) test statistic reject the null hypothesis that the taste distance coefficient in Column (4) is

[35] Due to perfect multicolinearity island, region and prefecture border effects can no longer be identified if island-, region- and prefecture-pair-specific fixed effects are introduced.

[36] In particular, we account for $4 + (4 \times 3)/2 = 10$ island-pair-specific fixed effects, $8 + (8 \times 7)/2 = 36$ region-pair-specific fixed effects and $46 + (46 \times 45)/2 = 1081$ prefecture-pair-specific fixed effects. Note that 32 prefecture-pair-specific fixed effects are dropped in Column (4) of Table 3 due to perfect multicoliniearity, which results whenever a prefecture possesses not more than a single wholesale market.

[37] A transportation infrastructure project, whose positive effects materialized predominantly at the island-pair level is the Great Seto Bridge, which since its opening in 1988 connects the adjacent islands of Honshû and Shikoku (cf. Konishi and Ono, 2024).

Table 3: *Accounting for unobserved heterogeneity (2SLS)*

Dependent variable: Relative price volatility for commodity c b/w the markets i and j ($REL\_PRICE\_VOL_{ijc}$)				
2nd-stage results				
Model:	**2SLS-FE**			
Specification:	**(1)**	**(2)**	**(3)**	**(4)**
$TASTE\_DIST_{ijc}$	1.0401**	0.9579**	1.2337**	1.2625
	(.4122)	(.4154)	(.4858)	(.8268)
$\ln DIST_{ij}$	0.0071***	0.0070***	0.0025	−0.0019
	(.0021)	(.0021)	(.0019)	(.0045)
$JOINT\_SUPPLIERS_{ijc}$	−0.0615***	−0.0617***	−0.0576***	−0.0574***
	(.0136)	(.0137)	(.0147)	(.0156)
$PREF\_BORD_{ij}$	−0.0104***	−0.0103***	−0.0067***	
	(.0022)	(.0022)	(.0022)	
$REGION\_BORD_{ij}$	−0.0021*	−0.0017		
	(.0012)	(.0013)		
$ISLAND\_BORD_{ij}$	0.0037			
	(.0028)			
Fixed effects:				
Market-specific fixed effects:	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓
Island-pair-specific fixed effects:	✗	✓	✓	✓
Region-pair-specific fixed effects:	✗	✗	✓	✓
Prefecture-pair-specific fixed effects:	✗	✗	✗	✓
Summary statistics:				
# of observations:	158,472	158,472	158,472	158,472
Kleibergen-Paap (p-value):	.0009	.0011	.0011	.0130
1st-stage results				
$HIST\_DIALECT\_SIM_{ij}$	−0.0149***	−0.0145***	−0.0140***	−0.0102**
	(.0037)	(.0037)	(.0035)	(.0037)
$\ln DIST_{ij}$	0.0018***	0.0018***	0.0024***	0.0043***
	(.0005)	(.0005)	(.0004)	(.0004)
$JOINT\_SUPPLIERS_{ijc}$	−0.0018	−0.0017	−0.0019	−0.0026
	(.0018)	(.0018)	(.0020)	(.0021)
$PREF\_BORD_{ij}$	0.0038***	0.0039***	0.0039***	
	(.0007)	(.0007)	(.0006)	
$REGION\_BORD_{ij}$	−0.0006	−0.0006		
	(.0006)	(.0006)		
$ISLAND\_BORD_{ij}$	0.0019*			
	(.0011)			
Fixed effects:				
Market-specific fixed effects:	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓
Island-pair-specific fixed effects:	✗	✓	✓	✓
Region-pair-specific fixed effects:	✗	✗	✓	✓
Prefecture-pair-specific fixed effects:	✗	✗	✗	✓
Summary statistics:				
# of observations:	158,472	158,472	158,472	158,472
R^2:	.2500	.2506	.2526	.2662
F-statistic (p-value):	.0003	.0004	.0004	.0108

Notes: Robust standard errors clustered at the commodity level; significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

statistically indistinguishable from zero at the 5% significance level. Reassuringly, the absolute size of point estimate (1.2625) is consistent with the magnitude of the estimates from the columns (1) to (3), rendering our baseline coefficient (1.0401) from Column (1) a rather conservative estimate for the effect of taste dissimilarity on relative price volatility.

To gauge the importance of a potential omitted variable bias for the 2SLS-versus-OLS gap in the estimates from Table 2, we now compare the 2SLS estimates from Columns (1) to (4) of Table 3 with the corresponding OLS estimates from the Columns (1) to (4) of Table A.10 (delegated to the Appendix). Neither in Table 3 nor in Table A.10, do we find statistically significant differences across the coefficients from the Columns (1) to (4). We therefore conclude that the 2SLS-versus-OLS gap in the estimates from the Tables 3 and A.10 can not be attributed to unobserved prefecture-pair-specific heterogeneity.

In summary, our results from Subsection 6.1 show that our baseline result from Table 2 is robust to the inclusion of island-, region- and prefecture-pair-specific fixed effects which account for unobserved bilateral heterogeneity. While these controls reduce the risk of omitted variable bias, they do not affect the magnitude of our OLS estimates. This suggests that the downward bias in the OLS estimates (relative to our 2SLS results) cannot be attributed to unobserved bilateral heterogeneity.

6.2 Assessing the plausibility of the exclusion restriction

The validity of our instrumental variable approach crucially depends on the exclusion restriction, according to which historical dialect similarity is not allowed to affect the volatility of LOP deviations through any other channel than through shaping the geography of regional tastes (see Subsection 5.1). While it is generally impossible to directly test the exclusion restriction, various studies (cf. Bound and Jaeger, 2000; Altonji et al., 2005; Nunn, 2008; Angrist et al., 2010; Basu et al., 2024; Levelu et al., 2024) have used zero-first-stage (ZFS) tests to provide suggestive evidence supporting the plausibility of the exclusion restriction.[38]

The ZFS test employs an auxiliary placebo regression to examine the reduced-form effect of the instrument on the outcome variable within a subsample for which it can be convincingly argued that the instrument's first-stage effect on the endogenous explanatory variable is zero. By focusing on this zero-first-stage subsample, the ZFS-test estimates the reduced-form effect of the instrument on the outcome variable without introducing omitted variable bias, which would otherwise arise from neglecting the endogenous explanatory variable. If the reduced-form coefficient of the instrument is approximately zero in this subsample, it provides suggestive evidence supporting the exclusion

[38] D'Haultfœuille et al. (2024) apply a similar argument in the context of the control function approach.

restriction.[39]

Since 97% of the respondents surveyed for the Linguistic Atlas of Japan were born between 1879 and 1903, the dialects they adopted in their youth provide a comprehensive reflection of the interactions that shaped the distribution of dialects and regional preferences up until the early 20th century in Japan. We therefore expect that our historical dialect similarity index correlates with the persistent distribution of localized preferences for fruits and vegetables that existed in pre-modern Japan (i.e., before the Meiji Restoration of 1868). However, it is unlikely to explain the distribution of varieties that were adopted and popularized at later stages.[40] Following this reasoning, the ZFS-test focuses on apples, which were introduced to Japan only in 1871 (cf. Sawamura et al., 1993) and did not become widely popular until after the end of World War II.[41]

Anecdotal evidence on the irrelevance of apples in early modern Japan comes from Nakagane (1873), who in its "World Customs" titled book from 1873 writes about Isaac Newton's Law of Universal Gravitation. To highlight Newton's eureka moment, the book features an illustration titled "Isaac Newton, discovering Gravity watching the fall" (see Figure A.2, delegated to the Appendix). In this illustration, Newton is depicted beneath a fruit tree, observing a fruit that has apparently just fallen. Upon closer inspection, it becomes clear that the fruit's shape and size do not resemble those of an apple. Instead, the author has replaced Newton's unfamiliar apple with a Japanese plum, which would have been more recognizable to his Japanese audience at the time.

Although apples are popular and widely available in Japan today, strong regional taste differences seem unlikely. The Japanese apple industry is highly geographically concentrated, with nearly all apples produced in Aomori Prefecture. Figures A.3 and A.4 (delegated to the Appendix) compare the geographic distribution of apple and nashi (Japanese pear) production. Unlike apples, nashis have been cultivated in Japan since the Asuka period (538‑710) (cf. Saito, 2016) and are grown in nearly every prefecture, primarily for local consumption. In contrast, apples are produced almost exclusively in Aomori Prefecture and, to a lesser extent, in Nagano Prefecture, from where they are distributed across the country.

Table 4 presents the results of the ZFS test, which regresses our instrument on the volatility of LOP deviations, using the same control variables as in Table 3. Notably, the effect of historical dialect similarity across Columns (1) to (4) in Table 4 is not only statistically insignificant but

[39]See van Kippersluis and Rietveld (2018) and Lal et al. (2024) for a more detailed exposition of the ZFS test.

[40]During the Edo period (1603-1868), Japan withdrew into a state of near autarky for over 200 years, significantly restricting the movement of goods, people, and ideas (see Bernhofen and Brown (2004, 2005) for an examination of the trade and welfare effects of Japan's unexpected transition from near autarky to nearly free trade after the end of its seclusion policy). Therefore, it is reasonable to assume that fruit and vegetable varieties not introduced to Japan before the country's complete seclusion in 1639 could not have been adopted until the policy ended in 1854 (marked by the Treaty of Kanagawa).

[41]Using production data from Shirai (2018), it can be shown that Japan's per capita apple production was below 0.5 kg until 1920. By 1950, production had increased more than tenfold to 4.9 kg per capita. As of 2020, the per capita production level had risen to 21.2 kg (cf. MAFF, 2023).

Table 4: *Zero-first-stage-test: apples*

Dependent variable: Relative price volatility of apples b/w the markets i and j ($REL\_PRICE\_VOL_{ij}$)				
Model:	**OLS-FE**			
Specification:	**(1)**	**(2)**	**(3)**	**(4)**
$HIST\_DIALECT\_SIM_{ij}$	-0.0022	-0.0020	-0.0017	-0.0068^{*}
	(.0026)	(.0027)	(.0026)	(.0037)
$\ln DIST_{ij}$	0.0043^{***}	0.0042^{***}	0.0029^{***}	0.0004
	(.0005)	(.0005)	(.0006)	(.0008)
$JOINT\_SUPPLIERS_{ij}$	0.0453^{***}	0.0454^{***}	0.0477^{***}	-0.0206^{**}
	(.0087)	(.0087)	(.0082)	(.0104)
$PREF\_BORD_{ij}$	-0.0039	-0.0043^{*}	0.0011	
	(.0024)	(.0024)	(.0022)	
$REGION\_BORD_{ij}$	0.0010	0.0015		
	(.0012)	(.0012)		
$ISLAND\_BORD_{ij}$	0.0024^{***}			
	(.0007)			
Fixed effects:				
Market-specific fixed effects:	✓	✓	✓	✓
Island-pair-specific fixed effects:	✗	✓	✓	✓
Region-pair-specific fixed effects:	✗	✗	✓	✓
Prefecture-pair-specific fixed effects:	✗	✗	✗	✓
Summary statistics:				
# of observations:	5, 112	5, 112	5, 112	5, 112
R^2:	0.9511	0.9513	0.9599	0.9846

Notes: Robust standard errors; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

also estimated to be very close to zero. This leads us to conclude that historical dialect similarity does not explain the relative price volatility between the markets in our zero-first-stage subsample, which is consistent with the exclusion restriction of our instrumental variable strategy. In Column (4), the negative effect of historical dialect similarity is somewhat larger in absolute size and marginally significant at the 10% level, which hints at a violation of the exclusion restriction when prefecture-pair-specific fixed effects are taken intoaccount.[42]

In their article van Kippersluis and Rietveld (2018) recommend to use the results of the ZFS-test as input for Conley et al.'s (2012) local-to-zero (LTZ) approach, which enables robust inference of instrumental variable estimates when the instrument is only plausibly exogenous. Conley et al.'s (2012) define plausible exogeneity by relaxing the IV exclusion restriction, allowing the instrument to have a small but nonzero effect on the outcome variable. An instrument is considered plausibly exogenous if prior information suggests that its effect on the outcome variable is near zero, but

[42]In Table A.11 (delegated to the Appendix), we provide further auxiliary evidence on the plausibility of our exclusion restriction. Using data from the Japanese Family Income and Expenditure Survey, we perform an alternative ZFS-test, showing that historical dialect similarity does not explain the volatility of relative hamburger prices across Japan. Our focus on hamburgers is inspired by The Economist's (2025) Big Mac index which uses burger prices to assess purchasing power parity across international borders, because the Big Mac is a highly standardized product and a symbol for cultural homogenization through globalization (cf. Levitt, 1983). Mc Donald's and its main competitor, MOS Burger, entered the Japanese market in 1971 and 1972, respectively. We therefore expect our historical instrument to be uncorrelated with the distribution of localized burger preferences in Japan.

not exactly zero. Following van Kippersluis and Rietveld's (2018) recommendation, we use the near-zero point estimates of historical dialect similarity's effect on relative price volatility from Table 4 to correct for potential violations of the exclusion restriction.[43]

Table 5: *Conley et al.'s (2012) LTZ results*

Dependent variable: Relative price volatility for commodity c b/w the markets i and j ($REL\_PRICE\_VOL_{ijc}$)				
Model:	**LTZ**			
Specification:	**(1)**	**(2)**	**(3)**	**(4)**
$TASTE\_DIST_{ijc}$	0.8929**	0.8203**	1.1123**	0.5953
	(.4101)	(.4136)	(.4831)	(.8220)
$\ln DIST_{ij}$	0.0074***	0.0072***	0.0028	0.0010
	(.0020)	(.0020)	(.0018)	(.0045)
$JOINT\_SUPPLIERS_{ijc}$	−0.0617***	−0.0619***	−0.0578***	−0.0592***
	(.0134)	(.0135)	(.0145)	(.0153)
$PREF\_BORD_{ij}$	−0.0099***	−0.0098***	−0.0062***	
	(.0021)	(.0021)	(.0022)	
$REGION\_BORD_{ij}$	−0.0022*	−0.0018		
	(.0012)	(.0013)		
$ISLAND\_BORD_{ij}$	0.0040			
	(.0027)			
Fixed effects:				
Market-specific fixed effects:	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓
Island-pair-specific fixed effects:	✗	✓	✓	✓
Region-pair-specific fixed effects:	✗	✗	✓	✓
Prefecture-pair-specific fixed effects:	✗	✗	✗	✓
Summary statistics:				
# of observations:	158,472	158,472	158,472	158,472

Notes: Robust standard errors clustered at the commodity level; significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5 replicates the results of Table 3 based on Conley et al.'s (2012) LTZ approach, using the point estimates from the ZFS-text in Table 4 to specify the prior for the effect of historical dialect similarity on the relative price volatility. Since these point estimates are both very close to zero and tightly estimated, it is unsurprising that the corrected IV estimates in Table 5 align well with our baseline results from Table 3. Adjusting for the the negative point estimates in Columns (1) to (4) of Table 4, intuitively reduces the absolute size of the coefficients that we obtain based on Conley et al.'s (2012) LTZ approach in Table 5 relative to their counterparts from Table 3.

In summary, the results from Subsection 6.2 provide strong support for the plausibility of our exclusion restriction while also demonstrating that our IV inference remains robust even if the

[43] Conley et al.'s (2012) propose four complementary inference strategies, which differ in terms of their informational requirements. According to van Kippersluis and Rietveld (2018) the LTZ approach constitutes an "elegant and user-friendly middle ground", which assumes that the prior for the effect of the instrument on the outcome variable follows a Normal distribution with known mean and variance. We take the point estimates from the first row of Table 4 as mean values and compute the corresponding variances as weighted averages, which are based on the standard errors from both the ZFS subsample and the remainder of the analysis sample (see van Kippersluis and Rietveld, 2018, pp. 319-20).

exclusion restriction is slightly violated.

6.3 Alternative instruments

To strengthen our analysis and address potential criticisms regarding the validity of our historical dialect instrument, we propose two complementary identification strategies. First, we refine our approach by focusing on a subsample of our data to construct a more narrowly defined instrument that closely reflects the historical taste distributions of three specific commodities. In a second step, we then complement our analysis by adding alternative supply-side instruments that exploit exogenous differences in climate conditions and soil quality.

Having introduced historical dialect similarity as a comprehensive summary measure reflecting the entirety of past interactions preserved in local dialects (see Subsection 5.1), raises the concern that our instrument is not only correlated with the taste distance from Eq.(12) but also with other outcomes such as intra-national migration and trade flows (cf. Falck et al., 2012; Bauernschuster et al., 2014; Lameli et al., 2015). To ensure that our instrument does not indirectly influence wholesale market integration by facilitating bilateral trade and migration, we construct a more narrowly defined historical dialect similarity measure. For this purpose, we focus on a single map of the Linguistic Atlas of Japan, which illustrates the geographic distribution of responses to the following question:

"In this area, when people normally say 'imo' (potato), which type of potato are they referring to (jagaimo [potato], satsumaimo [sweet potato], or satoimo [taro])? If someone says he/she wants to eat 'imo', which type of potato is he/she talking about?".[44]

The Linguistic Atlas of Japan (LAJ) documents four distinct responses to the above question, differentiating between potato (*jagaimo*, じゃが芋), sweet potato (*satsumaimo*, 薩摩芋), taro (*satoimo*, 里芋), and yam (*yamaimo*, 山芋). The recurring term *imo* (芋) in each name is the Japanese word for "tuber" or simply "potato".[45] Thus, when respondents were asked about their local potato variant, they likely named the crop that was most commonly consumed in their region at the time, thereby reflecting their region's historical taste preferences for these specific potato variants. By using the Jaccard index from Eq. (15) to compute the overlap in responses, we

[44]We reproduce LAJ map #179 in Figure A.5 (delegated to the Appendix). Since map #179 is not included in the LAJ Database (LAJDB) (cf. Kumagai, 2016), we digitized it using a GIS.

[45]Interestingly, each vegetable's Japanese name carries a geographic reference. Yam (*yamaimo*, 山芋) is native to Japan and literally means "mountain potato", reflecting its natural habitat and distinguishing it from taro (*satoimo*, 里芋), which translates to "village potato", underscoring taro's historical and cultural significance as a staple crop traditionally cultivated and consumed in rural areas. The sweet potato (*satsumaimo*, 薩摩芋) derives its name from Satsuma Province (now Kagoshima Prefecture in southern Japan), which played a crucial role in spreading the crop across Japan. Sweet potatoes were introduced to Satsuma in 1609 following the conquest of the Ryukyu Kingdom (modern-day Okinawa), where they had arrived from China in 1605 (cf. Laufer, 1929). The Japanese word for potato (*jagaimo*, じゃが芋), originates from "Jakarta potato", reflecting its introduction to Japan via Indonesia by Dutch traders in the early 17th century (cf. Laufer and Wilbur, 1938; Nunn and Qian, 2011).

therefore obtain a narrowly defined historical taste similarity measure, which is less likely to be correlated with unobserved market-pair specific variation than our more broadly defined historical similarity index from Eq. (16).[46]

Table 6: *Testing the localized tastes hypothesis based on a more narrow instrument*

Dependent variable: Relative price volatility for commodity c b/w the markets i and j ($REL\_PRICE\_VOL_{ijc}$)								
				2nd-stage results				
Model:	**OLS-FE**				**2SLS-FE**			
Specification:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
$TASTE\_DIST_{ijc}$	0.3569***	0.3571***	0.3235***	0.3437***	1.8268**	1.6017**	0.9605	−5.7483
	(.0241)	(.0265)	(.0269)	(.0294)	(.8301)	(.7032)	(.8406)	(6.6641)
$\ln DIST_{ij}$	0.0108***	0.0108***	0.0085***	0.0020	0.0042	0.0051*	0.0059*	0.0230
	(.0006)	(.0008)	(.0011)	(.0014)	(.0036)	(.0031)	(.0032)	(.0229)
$JOINT\_SUPPLIERS_{ijc}$	−0.0406***	−0.0407***	−0.0370***	−0.0334***	−0.0494***	−0.0485***	−0.0426***	0.0101
	(.0021)	(.0021)	(.0020)	(.0022)	(.0050)	(.0045)	(.0058)	(.0496)
$PREF\_BORD_{ij}$	−0.0097***	−0.0099***	−0.0041		−0.0076***	−0.0078***	−0.0038	
	(.0025)	(.0028)	(.0029)		(.0029)	(.0030)	(.0032)	
$REGION\_BORD_{ij}$	−0.0038***	−0.0038**			−0.0005	−0.0006		
	(.0014)	(.0015)			(.0021)	(.0021)		
$ISLAND\_BORD_{ij}$	0.0005				−0.0069			
	(.0014)				(.0043)			
Fixed effects:								
Market-specific fixed effects:	✓	✓	✓	✓	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓	✓	✓	✓	✓
Island-pair-specific fixed effects:	✗	✓	✓	✓	✗	✓	✓	✓
Region-pair-specific fixed effects:	✗	✗	✓	✓	✗	✗	✓	✓
Prefecture-pair-specific fixed effects:	✗	✗	✗	✓	✗	✗	✗	✓
Summary statistics:								
# of observations:	15,168	15,168	15,168	15,168	13,686	13,686	13,686	13,686
R^2:	.8436	.8436	.8466	.8569	-	-	-	-
Kleibergen-Paap (p-value):	-	-	-	-	.0001	.0000	.0002	.2517
				1st-stage results				
$HIST\_DIALECT\_SIM_{ij}$					−0.0030***	−0.0034***	−0.0030***	−0.0013
					(.0008)	(.0008)	(.0008)	(.0011)
$\ln DIST_{ij}$					0.0039***	0.0039***	0.0033***	0.0035***
					(.0002)	(.0002)	(.0003)	(.0004)
$JOINT\_SUPPLIERS_{ijc}$					0.0056***	0.0058***	0.0064***	0.0074***
					(.0008)	(.0008)	(.0008)	(.0009)
$PREF\_BORD_{ij}$					−0.0002	−0.0003	0.0011*	
					(.0007)	(.0007)	(.0007)	
$REGION\_BORD_{ij}$					−0.0017***	−0.0020***		
					(.0005)	(.0005)		
$ISLAND\_BORD_{ij}$					0.0049***			
					(.0005)			
Fixed effects:								
Market-specific fixed effects:					✓	✓	✓	✓
Commodity-specific fixed effects:					✓	✓	✓	✓
Island-pair-specific fixed effects:					✗	✓	✓	✓
Region-pair-specific fixed effects:					✗	✗	✓	✓
Prefecture-pair-specific fixed effects:					✗	✗	✗	✓
Summary statistics:								
# of observations:					13,686	13,686	13,686	13,686
R^2:					.3452	.3470	.3588	.4398
F-statistic (p-value):					.0001	.0000	.0002	.2728

Notes: Robust standard errors; significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In Table 6, we reexamine the effect of taste heterogeneity on the relative price volatility of potatoes, sweet potatoes, and taro. Unlike in Column (4) of Table 2, the OLS result in Column (1) of Table 6 is statistically different from zero and does not change across the Columns (2) to (4) even when controlling for unobserved bilateral heterogeneity through island-, region-,

[46]To minimize data loss due to missing responses, we use a 50-km radius around wholesale market locations to match responses from the Linguistic Atlas of Japan with our market data.

and prefecture-pair-specific fixed effects. However, while statistically different from zero, these estimates are substantially smaller than the corresponding 2SLS results in Table 3, which we attribute to attenuation bias caused by measurement error.

To account for a possible downward bias affecting our OLS estimates, we use our more narrowly defined historical dialect/taste similarity index to instrument the commodity-specific taste distance for potatoes, sweet potatoes, and taro in Columns (5) to (8). In Columns (5) to (7) we have a strong 1^{st} stage with the expected negative signs and somewhat smaller point estimates as in Table 3, which turns insignificant once a large number of prefecture-pair-specific fixed effects is taken into account in Column (8). Our 2^{nd}-stage results in Columns (5) and (6) indicate that taste differences for potatoes, sweet potatoes, and taro have a statistically significant and economically meaningful impact on relative price volatility, which is somewhat more pronounced than in Table 3. Reassuringly, we find that the OLS-versus-2SLS gaps in the parameter estimates are comparable to those from Tables 3 and A.10. However, as we introduce increasingly demanding sets of bilateral fixed effects, the power of our instrument diminishes. As a consequence the estimate in Column (8) suffers from a weak instrument problem, with a 1^{st}-stage F-statistic of just 1.2, leading to an unreliable estimation result and substantially inflated standard errors.

By replicating our baseline results using a more narrowly defined instrument that directly links historical taste to geography, Table 6 not only reaffirms the magnitude and significance of our preferred estimates but also strengthens the credibility of our exclusion restriction.

So far our IV approach has focused on how persistent, inter-regional taste differences have been shaped by past cultural exchange. As a complimentary source for exogenous variation, we now utilize differences in agro-climatic endowments across Japanese regions to explain how persistent food preferences have been influenced by the relative abundance of certain crops (cf. Atkin, 2013; Head and Mayer, 2013). To this end, we extend our dataset by incorporating geospatial data provided by Japan's Ministry of Land, Infrastructure, Transport and Tourism (MLIT) to account for differences in climate conditions and soil quality. Specifically, we use the Climate Mesh Data from MLIT's National Land Planning Bureau, which offers detailed 1km $\times$ 1km grid-level information on (*i.*) annual precipitation, (*ii.*) annual temperature, and (*iii.*) total annual sunshine hours. From this, we compute market-pair-specific differences by averaging all observations within a 50-km-radius circle radius around each market location. Additionally, information on soil quality is drawn from MLIT's Land Classification Mesh Data, which provides 1km $\times$ 1km grid-level details on (*i.*) surface geology, (*ii.*) landform classification and (*iii.*) soil type. As each category contains various subcategories, we calculate market-pair-specific differences in two steps: first, we determine the share of each subcategory within a 50-km radius around each market location. Next, we compute the Manhattan distance over bilateral shares to derive three market-pair-specific measures that capture average differences in surface geology, landform classification, and soil type.

In Table 7, we extend our instrumental variable analysis by incorporating additional instruments that leverage exogenous variations in climate conditions and soil quality as 1st-stage predictors. As a benchmark, Column (1) reproduces our baseline estimates from Column (5) of Table 2. Columns (2) and (3) introduce climate- and soil-based instruments individually, while Column (4) combines both sets of instruments. Finally, Column (5) presents results using the combined climate- and soil-based instruments while omitting the historical dialect similarity index as a 1st-stage instrument.

Reassuringly, we find that the magnitude and significance of the 1st-stage effect of historical dialect similarity are robust against the inclusion of additional climate- and soil-based instruments. Hansen's J statistic indicates that we cannot reject the joint validity of our instruments in Columns (2) and (3). However, when all instruments are combined, joint validity is rejected at the 10% significance level. The climate-based instruments in Columns (2), (4), and (5) exhibit the expected positive coefficients and are statistically significant in all but one case. Regarding the soil-based instruments, differences in surface geology and landform classification show a negative and significant effect on taste heterogeneity in Columns (3), (4), and (5). In contrast, differences in soil types positively affect commodity-specific taste differences, though the estimate in Column (4) is imprecise.

Across all columns of Table 7, we observe a positive and significant effect of taste heterogeneity on the volatility of LOP deviations, which becomes slightly larger when differences in exogenous climate conditions are introduced as additional 1st-stage instruments. Our 2SLS results indicate a weak-instrument problem (with the IV's relative bias slightly exceeding Stock and Yogo's (2005) 10% threshold) when relying on multiple climate- and soil-based instruments. Reassuringly, the weak-instrument-robust test statistics from Anderson and Rubin (1949) and Stock and Wright (2000) suggest that we can reject the null hypothesis that the taste distance coefficient in Columns (2) through (5) is statistically indistinguishable from zero at the 1% significance level. Despite this re-affirmative result, it is worth noting that the p value of the 1st-stage F statistic increases by factor three when the historical dialect similarity index is excluded as an instrument in Column (5).

By relying on a more narrow definition of our historical dialect instrument and by leveraging exogenous variation in climate conditions and soil quality to construct complementary instruments, we provide robust evidence on the validity and importance of historical dialect similarity as our preferred instrument to predict persistent inter-regional taste differences that have historically developed whenever regions were linked through cultural exchange.

Table 7: *Alternative Instruments*

Dependent variable: Relative price volatility for commodity c b/w the markets i and j ($REL\_PRICE\_VOL_{ijc}$)					
			2nd-stage results		
Model:	**2SLS-FE**				
Specification:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$TASTE\_DIST_{ijc}$	1.0401**	1.8236***	0.8660**	1.6834***	2.3185***
	(.4122)	(.5051)	(.3887)	(.4830)	(.8004)
$\ln DIST_{ij}$	0.0071***	0.0052**	0.0075***	0.0056**	0.0040
	(.0021)	(.0023)	(.0021)	(.0022)	(.0028)
$JOINT\_SUPPLIERS_{ijc}$	−0.0615***	−0.0601***	−0.0618***	−0.0603***	−0.0592***
	(.0136)	(.0142)	(.0136)	(.0141)	(.0149)
$PREF\_BORD_{ij}$	−0.0104***	−0.0149***	−0.0094***	−0.0141***	−0.0178***
	(.0022)	(.0028)	(.0020)	(.0027)	(.0050)
$REGION\_BORD_{ij}$	−0.0021*	−0.0020	−0.0021*	−0.0020	−0.0019
	(.0012)	(.0015)	(.0012)	(.0014)	(.0017)
$ISLAND\_BORD_{ij}$	0.0037	0.0025	0.0040	0.0027	0.0017
	(.0028)	(.0032)	(.0027)	(.0031)	(.0036)
Fixed effects:					
Market-specific fixed effects:	✓	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓	✓
Summary statistics:					
# of observations:	158,472	158,472	158,472	158,472	158,472
Kleibergen-Paap (p-value):	.0009	.0052	.0011	.0300	.0379
Hansen J statistic (p-value):	–	.2082	.2265	.0504	.1469
			1st-stage results		
$HIST\_DIALECT\_SIM_{ij}$	−0.0149***	−0.0140***	−0.0141***	−0.0136***	
	(.0006)	(.0006)	(.0006)	(.0006)	
$RAIN\_DIFF_{ij}$		0.0013***		0.0011***	0.0016***
		(.0004)		(.0004)	(.0004)
$TEMP\_DIFF_{ij}$		0.0055***		0.0053***	0.0059***
		(.0005)		(.0005)	(.0005)
SUN_DIFF_{ij}		0.0011**		0.0010**	0.0005
		(.0004)		(.0004)	(.0004)
$SURFAC\_DIFF_{ij}$			−0.0628***	−0.0673***	−0.0621***
			(.0161)	(.0162)	(.0162)
$LANDFORM\_DIFF_{ij}$			−0.0745***	−0.0307*	−0.0823***
			(.0174)	(.0177)	(.0175)
$SOIL\_DIFF_{ij}$			0.0951***	0.0358	0.0654***
			(.0244)	(.0251)	(.0252)
$\ln DIST_{ij}$	0.0018***	0.0016***	0.0018***	0.0017***	0.0022***
	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)
$JOINT\_SUPPLIERS_{ijc}$	−0.0018***	−0.0020***	−0.0019***	−0.0020***	−0.0020***
	(.0003)	(.0003)	(.0003)	(.0003)	(.0003)
$PREF\_BORD_{ij}$	0.0038***	0.0037***	0.0035***	0.0035***	0.0051***
	(.0003)	(.0003)	(.0003)	(.0003)	(.0003)
$REGION\_BORD_{ij}$	−0.0006***	−0.0009***	−0.0007***	−0.0010***	−0.0007***
	(.0002)	(.0002)	(.0002)	(.0002)	(.0002)
$ISLAND\_BORD_{ij}$	0.0019***	0.0019***	0.0019***	0.0019***	0.0017***
	(.0002)	(.0002)	(.0002)	(.0002)	(.0002)
Fixed effects:					
Market-specific fixed effects:	✓	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓	✓
Summary statistics:					
# of observations:	158,472	158,472	158,472	158,472	158,472
R^2:	.2500	.2511	.2503	.2513	.2495
F-statistic (p-value):	.0003	.0005	.0014	.0022	.0064

Notes: Robust standard errors clustered at the commodity level; significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

7 Conclusion

Historically determined differences in regional taste turn out to have a long-lasting effect on the volatility of relative market prices. Using wholesale market data on fruits and vegetables from Japan, we show that bilateral commodity-specific taste differences – instrument by historical dialect similarity – explain approximately 9% of of the mean volatility in law-of-one-price deviations at the commodity level. Attenuation bias due to measurement error poses a key challenge in identifying the causal effect of taste heterogeneity on the volatility of law-of-one-price deviations, as localized tastes cannot be directly observed and must be estimated, inevitably introducing some margin of error.

To account for measurement error and other potential sources of endogeneity, we propose a novel instrument, which relies on variation in historical dialects to predict persistent inter-regional differences in historically determined tastes. To strengthen the validity of our instrument, we provide additional evidence supporting the exclusion restriction. Specifically, using a subsample for which the taste channel is arguably muted, we conduct a placebo regression, confirming that historical dialect similarity has a negligible impact on the volatility of relative market prices. Focusing on a subsample of our fruits and vegetable data, for which it can be convincingly argued that the taste channel is muted, we perform a placebo regression, which confirms that historical dialect similarity has a negligible impact on the volatility of relative market prices. Alternatively, we also confirm our baseline results based on a more narrowly defined instrument, which relies on a specific question from the Linguistic Atlas of Japan, to directly measure the spatial distribution of historical food preferences in Japan.

By providing robust causal evidence on the effect of taste heterogeneity – complementing recent research on the market-integrating impact of reduced information frictions – we also want to emphasis a key distinction between these two effects. While eliminating information frictions is expected to yield gains from market integration, no such gains arise if the observed lack of market integration simply reflects underlying differences in consumer preferences. To gauge the potential for future market integration it therefore is crucial to quantify the long-lasting effect of regional taste heterogeneity.

References

Aizenman, J. and E. Brooks (2008): "Globalization and Taste Convergence: The Cases of Wine and Beer," *Review of International Economics*, 16, 217–33. 17

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "An Evaluation of Instrumental Variable

Strategies for Estimating the Effects of Catholic Schooling," *Journal of Human Resources*, 40, 791–821. 4, 25

Anderson, J. E. and E. van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170–92. 1

Anderson, J. E. and P. Zhang (2025): "Latent Exports: Almost Ideal Gravity and Zeros," *The Review of Economics and Statistics*, 107, 221–39. 12

Anderson, T. W. and H. Rubin (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, 20, 46–63. 5, 23, 32

Angrist, J., V. Lavy, and A. Schlosser (2010): "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," *Journal of Labor Economics*, 28, 773–824. 4, 25

Atkin, D. (2013): "Trade, Tastes, and Nutrition in India," *American Economic Review*, 103, 1629–63. 1, 3, 5, 12, 13, 14, 16, 17, 31

——— (2016): "The Caloric Costs of Culture: Evidence from Indian Migrants," *American Economic Review*, 106, 1144–81. 1, 16, 17

Atkin, D., E. Colson-Sihra, and M. Shayo (2021): "How Do We Choose Our Identity? A Revealed Preference Approach Using Food Consumption," *Journal of Political Economy*, 129, 1193–251. 14

Aw, B. Y., Y. Lee, and H. Vandenbussche (2023): "Quantifying Consumer Taste in Trade: Evidence from the Food Industry," CESifo Working Paper Series 10234, CESifo. 5, 19

Bailey, M., A. Gupta, S. Hillenbrand, T. Kuchler, R. Richmond, and J. Stroebel (2021): "International Trade and Social Connectedness," *Journal of International Economics*, 129, 103418. 1

Bargain, O. (2024): "Globalization and Cultural Spillover in Trade: Evidence from the Japanese Food Culture," *Review of World Economics*, 160, 55–73. 17

Basu, A. K., T. Byambasuren, N. H. Chau, and N. Khanna (2024): "Cooking Fuel Choice and Child Mortality in India," *Journal of Economic Behavior & Organization*, 222, 240–65. 25

Bauernschuster, S., O. Falck, S. Heblich, J. Suedekum, and A. Lameli (2014): "Why Are Educated and Risk-loving Persons more Mobile across Regions?" *Journal of Economic Behavior & Organization*, 98, 56–69. 4, 20, 29

Beltramello, A., K. D. Backer, V. Mercader, and L. Moussiegt (2011): “Opening Japan: Comparisons with Other G20 Countries and Lessons Learned from International Experience,” OECD Science, Technology and Industry Working Papers 2011/02, OECD Publishing, Paris. 3

Bernhofen, D. M. and J. C. Brown (2004): “A Direct Test of the Theory of Comparative Advantage: The Case of Japan,” *Journal of Political Economy*, 112, 48–67. 3, 26

——— (2005): “An Empirical Assessment of the Comparative Advantage Gains from Trade: Evidence from Japan,” *American Economic Review*, 95, 208–25. 4, 26

Bisin, A. and T. Verdier (2001): “The Economics of Cultural Transmission and the Dynamics of Preferences,” *Journal of Economic Theory*, 97, 298–319. 1, 2, 17

——— (2014): “Chapter 17 - Trade and Cultural Diversity,” in *Handbook of the Economics of Art and Culture*, ed. by V. A. Ginsburgh and D. Throsby, Elsevier, vol. 2 of *Handbook of the Economics of Art and Culture*, 439–84. 16, 17

Bound, J. and D. A. Jaeger (2000): “Do Compulsory School Attendance Laws alone Explain the Association between Quarter of Birth and Earnings?” in *Research in Labor Economics*, vol. 19, 83–108. 25

Bove, V. and G. Gokmen (2018): “Genetic Distance, Trade, and the Diffusion of Development,” *Journal of Applied Econometrics*, 33, 617–23. 17

Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012): “The Evolution of Brand Preferences: Evidence from Consumer Migration,” *American Economic Review*, 102, 2472-2508. 1, 17

Ceglowski, J. (2003): “The Law of One Price: Intranational Evidence for Canada,” *Canadian Journal of Economics*, 36, 373–400. 1

Chen, N. (2004): “Intra-national versus International Trade in the European Union: Why Do National Borders Matter?” *Journal of International Economics*, 63, 93–118. 1

Cohen, L., U. G. Gurun, and C. Malloy (2017): “Resident Networks and Corporate Connections: Evidence from World War II Internment Camps,” *The Journal of Finance*, 72, 207–48. 1

Colson-Sihra, E., J. D. Sousa, and T. Mayer (2023): “Tastes, Geography and Culture,” mimeo, Hebrew University of Jerusalem. 1, 2, 3, 12, 13, 14, 15, 16, 19, 21

Combes, P.-P., M. Lafourcade, and T. Mayer (2005): "The Trade-creating Effects of Business and Social Networks: Evidence from France," *Journal of International Economics*, 66, 1–29. 1

Conley, T. G., C. B. Hansen, and P. E. Rossi (2012): "Plausibly Exogenous," *The Review of Economics and Statistics*, 94, 260–72. 4, 27, 28

Crucini, M. J., M. Shintani, and T. Tsuruga (2010): "The Law of One Price without the Border: The Role of Distance versus Sticky Prices," *The Economic Journal*, 120, 462–80. 1, 2, 9

——— (2015): "Noisy Information, Distance and Law of One Price Dynamics across US Cities," *Journal of Monetary Economics*, 74, 52–66. 1

Deaton, A. S. and J. Muellbauer (1980): "An Almost Ideal Demand System," *American Economic Review*, 70, 312–26. 3, 12, 14

D'Haultfœuille, X., S. Hoderlein, and Y. Sasaki (2024): "Testing and Relaxing the Exclusion Restriction in the Control Function Approach," *Journal of Econometrics*, 240, 105075. 25

Disdier, A.-C. and K. Head (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade," *The Review of Economics and Statistics*, 90, 37–48. 1

Disdier, A.-C. and T. Mayer (2007): "Je t'aime, moi non plus: Bilateral Opinions and International trade," *European Journal of Political Economy*, 23, 1140–59. 17

Dunlevy, J. A. (2006): "The Influence of Corruption and Language on the Protrade Effect of Immigrants: Evidence from the American States," *The Review of Economics and Statistics*, 88, 182–86. 1

Elberg, A. (2016): "Sticky Prices and Deviations from the Law of One Price: Evidence from Mexican Micro-price Data," *Journal of International Economics*, 98, 191–203. 1

Engel, C. and J. H. Rogers (1996): "How Wide Is the Border?" *The American Economic Review*, 86, 1112–25. 1, 2, 9, 11, 12

Fajgelbaum, P. D. and A. K. Khandelwal (2016): "Measuring the Unequal Gains from Trade," *The Quarterly Journal of Economics*, 131, 1113–80. 12

Falck, O., S. Heblich, A. Lameli, and J. Südekum (2012): "Dialects, Cultural Identity, and Economic Exchange," *Journal of Urban Economics*, 72, 225–39. 2, 4, 15, 17, 18, 19, 20, 29

Feenstra, R. and C. Hong (2022): "China's Import Demand for Agricultural Products: The Impact of the Phase One Trade Agreement," *Review of International Economics*, 30, 345–68. 12

Felbermayr, G. and J. Gröschl (2014): "Within U.S. Trade And The Long Shadow Of The American Secession," *Economic Inquiry*, 52, 382–404. 1

Felbermayr, G. J. and F. Toubal (2010): "Cultural Proximity and Trade," *European Economic Review*, 54, 279–93. 17

Fuller, W. A. (1987): *Measurement Error Models*, John Wiley & Sons, Ltd. 19

Grossman, G. M. (1998): "Comment: Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" in *The Regionalization of the World Economy*, ed. by J. A. Frankel, Chicago: The University of Chicago Press, 29–32. 1

Guiso, L., P. Sapienza, and L. Zingales (2009): "Cultural Biases in Economic Exchange?" *The Quarterly Journal of Economics*, 124, 1095–31. 17, 20

Hausman, J. (2001): "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left," *The Journal of Economic Perspectives*, 15, 57–67. 19

Head, K. and T. Mayer (2013): "What Separates Us? Sources of Resistance to Globalization," *The Canadian Journal of Economics*, 46, 1196–231. 1, 2, 3, 16, 17, 20, 21, 31

——— (2014): "Chapter 3 - Gravity Equations: Workhorse, Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by K. R. Elhanan Helpman and G. Gopinath, Elsevier, vol. 4 of *Handbook of International Economics*, 131–95. 1, 9, 11

——— (2021): "The United States of Europe: A Gravity Model Evaluation of the Four Freedoms," *Journal of Economic Perspectives*, 35, 23–48. 2, 9

Head, K., T. Mayer, and J. Ries (2010): "The Erosion of Colonial Trade Linkages after Independence," *Journal of International Economics*, 81, 1–14. 17

Hellmanzik, C. and M. Schmitz (2015): "Virtual Proximity and Audiovisual Services Trade," *European Economic Review*, 77, 82–101. 17

Herander, M. G. and L. A. Saavedra (2005): "Exports and the Structure of Immigrant-Based Networks: The Role of Geographic Proximity," *The Review of Economics and Statistics*, 87, 323–35. 1

Hillberry, R. and D. Hummels (2003): "Intranational Home Bias: Some Explanations," *The Review of Economics and Statistics*, 85, 1089–92. 1

Hillrichs, D. and G. Vannoorenberghe (2022): "Trade Costs, Home Bias and the Unequal Gains from Trade," *Journal of International Economics*, 139, 103684. 12

Hinskens, F., P. Auer, and P. Kerswill (2005): "The Study of Dialect Convergence and Divergence: Conceptual and Methodological Considerations," in *Dialect Change: Convergence and Divergence in European Languages*, ed. by P. Auer, F. Hinskens, and P. Kerswill, Cambridge University Press, 1–48. 2, 18

Hortaçsu, A., F. A. Martínez-Jerez, and J. Douglas (2009): "The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre," *American Economic Journal: Microeconomics*, 1, 53–74. 1

Huang, R. R. (2007): "Distance and Trade: Disentangling Unfamiliarity Effects and Transport Cost Effects," *European Economic Review*, 51, 161–81. 1

Hummels, D. and K. Y. Lee (2018): "The Income Elasticity of Import Demand: Micro Evidence and an Application," *Journal of International Economics*, 113, 20–34. 12

Hut, S. (2020): "Determinants of Dietary Choice in the US: Evidence from Consumer Migration," *Journal of Health Economics*, 72, 102327. 17

Hut, S. and E. Oster (2022): "Changes in Household Diet: Determinants and Predictability," *Journal of Public Economics*, 208, 104620. 17

Jaccard, P. (1901): "Étude comparative de la distribution florale dans une portion des Alpes et des Jura," *Bulletin del la Société Vaudoise des Sciences Naturelles*, 37, 547–79. 18

JCDP (2025): "Japan-Asia Climate Date Program: Japan Climate Atlas (1991-2020)," Available at: `https://jcdp.jp/japan/` (Accessed: 2025-02-08). 5

Jeszenszky, P., Y. Hikosaka, S. Imamura, and K. Yano (2019): "Japanese Lexical Variation Explained by Spatial Contact Patterns," *ISPRS International Journal of Geo-Information*, 8(9), 1–30. 16, 19

Jiang, W. (2017): "Have Instrumental Variables Brought Us Closer to the Truth," *The Review of Corporate Finance Studies*, 6, 127–40. 19

Kano, K., T. Kano, and K. Takechi (2013): "Exaggerated Death of Distance: Revisiting Distance Effects on Regional Price Dispersions," *Journal of International Economics*, 90, 403–13. 6, 9

Kimura, A. H. (2021): "Pickles and Agrobiodiversity: A Foodway and Traditional Vegetable Varieties in Japan," *Agriculture and Human Values*, 38, 1079–96. 17

Kleibergen, F. and R. Paap (2006): "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133, 97–126. 21

Konishi, Y. and A. Ono (2024): "Do Winners Win More from Transport Megaprojects? Evidence from the Great Seto Bridge in Japan," Keio-IES Discussion Paper Series 2024-018, Institute for Economics Studies, Keio University. 23

Kumagai, Y. (2016): "Chapter 20: Developing the Linguistic Atlas of Japan Database and Advancing Analysis of Geographical Distributions of Dialects," in *The Future of Dialects*, ed. by M.-H. Côté, R. Knooihuizen, and J. Nerbonne, Berlin: Language Science Press, 333–62. 2, 7, 29

Kónya, I. and H. Ohashi (2007): "International Consumption Patterns among High-income Countries: Evidence from the OECD Data," *Review of International Economics*, 15, 744–57. 17

Lal, A., M. Lockhart, Y. Xu, and Z. Zu (2024): "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice Based on 67 Replicated Studies," *Political Analysis*, 32, 521–40. 19, 26

Lameli, A., V. Nitsch, J. Südekum, and N. Wolf (2015): "Same Same But Different: Dialects and Trade," *German Economic Review*, 16, 290–306. 4, 17, 20, 29

Laufer, B. (1929): "The American Plant Migration," *The Scientific Monthly*, 28, 239–51. 29

Laufer, B. and C. M. Wilbur (1938): "The American Plant Migration: Part I: The Potato," *Publications of the Field Museum of Natural History. Anthropological Series*, 28, 1–132. 29

Lawrence, R. Z. (1987): "Imports in Japan: Closed Markets," *Brookings Papers on Economic Activity*, 18, 517–54. 3

——— (1991): "How Open is Japan?" in *Trade with Japan: Has the Door Opened Wider?*, ed. by P. Krugman, Chicago: University of Chicago Press, 9–50. 3

Lendle, A., M. Olarreaga, S. Schropp, and P.-L. Vézina (2016): "There Goes Gravity: Ebay and the Death of Distance," *The Economic Journal*, 126, 406–41. 1

Levelu, A., A. M. Mayda, and G. Orefice (2024): "Deep Trade Agreements and International Migration: The Role of Visa Provisions," *Journal of Economic Geography*, 25, 75–92. 25

Levitt, T. (1983): "The Globalization of Markets," *Harvard Business Review*, 61, 92–102. 27

MAFF (2023): "The 96th Statistical Yearbook of Ministry of Agriculture, Forestry and Fisheries," Available at: https://www.maff.go.jp/e/data/stat/96th/index.html (Accessed: 2025-03-09). 26

——— (2025a): "Pickled," Available at: https://www.maff.go.jp/e/policies/market/dento_syoku/bunrui/tukemono.html (Accessed: 2025-02-08). 3, 17

——— (2025b): "Traditional Foods in Japan," Available at: https://www.maff.go.jp/e/policies/market/dento_syoku/index.html (Accessed: 2025-02-08). 16, 17

Mason, M. M. (2012): *Dominant Narratives of Colonial Hokkaido and Imperial Japan: Envisioning the Periphery and the Modern Nation-State*, New York: Palgrave Macmillan. 19

Maystre, N., J. Olivier, M. Thoenig, and T. Verdier (2014): "Product-based Cultural Change: Is the Village Global?" *Journal of International Economics*, 92, 212–30. 16, 17

McCallum, J. (1995): "National Borders Matter: Canada-U.S. Regional Trade Patterns," *The American Economic Review*, 85, 615–23. 1

Melitz, J. and F. Toubal (2014): "Native Language, Spoken Language, Translation and Trade," *Journal of International Economics*, 93, 351–63. 17

Millimet, D. L. and T. Osang (2007): "Do State Borders Matter for U.S. Intranational Trade? The Role of History and Internal Migration," *Canadian Journal of Economics*, 40, 93–126. 1

Möhlmann, L. J., S. Ederveen, H. L. F. De Groot, and G. J. M. Linders (2010): "Intangible Barriers to International Trade: A Sectoral Approach," in *The Gravity Model in International Trade: Advances and Applications*, ed. by P. A. G. van Bergeijk and S. Brakman, Cambride: Cambridge University Press, 224‐252. 17

Nakagane, M. (1873): *World Customs (sekai fūzoku ôrai gaihen)*, Tokyo: Izumiya Ichibee. 26, 58

National Language Research Institute (1966): *Introduction to the Linguistic Atlas of Japan (nihon gengo chizu kaisetsu hôhô)*, Tokyo: Printing Bureau, Ministry of Finance. 2, 7, 8

Nevo, A. (2011): "Empirical Models of Consumer Behavior," *Annual Review of Economics*, 3, 51–75. 12, 13, 14

Nitsch, V. and N. Wolf (2013): "Tear Down this Wall: On the Persistence of Borders in Trade," *Canadian Journal of Economics*, 46, 154–79. 1

Nunn, N. (2008): "The Long-term Effects of Africa's Slave Trades," *The Quarterly Journal of Economics*, 123, 139–76. 4, 25

Nunn, N. and N. Qian (2011): "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment," *The Quarterly Journal of Economics*, 126, 593–650. 29

Ochsner, C. and F. Roesel (2020): "Migrating Extremists," *The Economic Journal*, 130, 1135–72. 17

OECD (2024): "Recruiting Immigrant Workers: Japan 2024," OECD Publishing, Paris. 3

Onishi, T. (2011): "16. Mapping the Japanese language," in *Volume 2 Language Mapping*, ed. by A. Lameli, R. Kehrein, and S. Rabanus, Berlin, New York: De Gruyter Mouton, 333–60. 8, 19

Oster, E. (2018): "Diabetes and Diet: Purchasing Behavior Change in Response to Health Information," *American Economic Journal: Applied Economics*, 10, 308-48. 17

Pancost, A. and G. Schaller (2024): "Measuring Measurement Error," mimeo, University of Texas at Austin. 19

Parsley, D. C. and S.-J. Wei (1996): "Convergence to the Law of One Price Without Trade Barriers or Currency Fluctuations," *The Quarterly Journal of Economics*, 111, 1211–36. 1, 2

——— (2001): "Explaining the Border Effect: The Role of Exchange Rate Variability, Shipping Costs, and Geography," *Journal of International Economics*, 55, 87–105. 1

Parsons, C. and P.-L. Vézina (2018): "Migrant Networks and Trade: The Vietnamese Boat People as a Natural Experiment," *The Economic Journal*, 128, F210–F234. 1

Rauch, J. E. (1999): "Networks versus Markets in International Trade," *Journal of International Economics*, 48, 7–35. 1

Rauch, J. E. and V. Trindade (2002): "Ethnic Chinese Networks in International Trade," *The Review of Economics and Statistics*, 84, 116–30. 1, 17

Saito, T. (2016): "Advances in Japanese Pear Breeding in Japan," *Breeding Science*, 66, 46–59. 26

Sawamura, K., Y. Harada, T. Kikuchi, and J. M. Ogawa (1993): "The Apple Industry in Japan: A Historical Sketch and Diseases Specific to the Region," *Plant Disease*, 77, 83–124. 26

Saxonhouse, G. R. (1993): "What Does Japanese Trade Structure Tell Us about Japanese Trade Policy?" *Journal of Economic Perspectives*, 7, 21–43. 3

Schmidt, J. E. (2010): "12. Language and Space: The Linguistic Dynamics Approach," in *Volume 1 Theories and Methods, An International Handbook of Linguistic Variation*, ed. by P. Auer and J. E. Schmidt, Berlin, New York: De Gruyter Mouton, 201–25. 2, 17

Shigeaki, S. (1991): "Outline History of the Colonisation of Hokkaido, 1870-1930," in *Migrants in Agricultural Development: A Study of Intrarural Migration*, ed. by J. A. Mollett, London: Palgrave Macmillan, 92–110. 19

Shirai, I. (2018): "The Formation and Maintenance of a Regional Brand: A Case Study of Apple Production in Aomori Prefecture in Modern Japan," *Japanese Research in Business History*, 35, 83–124. 26

Staehle, H. (1934): "The Reaction of Consumers to Changes in Prices and Income: A Quantitative Study in Immigrants' Behavior," *Econometrica*, 2, 59–72. 17

Steinwender, C. (2018): "Real Effects of Information Frictions: When the States and the Kingdom Became United," *American Economic Review*, 108, 657–96. 1

Stock, J. H. and J. H. Wright (2000): "GMM with Weak Identification," *Econometrica*, 68, 1055–96. 5, 23, 32

Stock, J. H. and M. Yogo (2005): "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. by D. W. K. Andrews and J. H. Stock, Cambridge: Cambridge University Press, 80–108. 23, 32

Subramanian, S. and A. Deaton (1996): "The Demand for Food and Calories," *Journal of Political Economy*, 104, 133–62. 14

The Economist (2025): "Our Big Mac Index Shows How Burger Prices Differ Across Borders," Available at: `https://www.economist.com/interactive/big-mac-index` (Accessed: 2025-03-09). 27

UN (2020): "National Accounts Statistics: Main Aggregates and Detailed Tables, 2020," United Nations, Department of Economic and Social Affairs. 3

UNESCO (2013): "Decision of the Intergovernmental Committee: 8.COM 8.17," Available at: `https://ich.unesco.org/en/decisions/8.COM/8.17` (Accessed: 2025-02-08). 3, 16

van Kippersluis, H. and C. A. Rietveld (2018): "Beyond Plausibly Exogenous," *The Econometrics Journal*, 21, 316–31. 4, 26, 27, 28

Wolf, H. C. (2000): "Intranational Home Bias in Trade," *The Review of Economics and Statistics*, 82, 555–63. 1

Yilmazkuday, H. (2012): "Understanding Interstate Trade Patterns," *Journal of International Economics*, 86, 158–66. 1

Yotov, Y. V. (2012): "A Simple Solution to the Distance Puzzle in International Trade," *Economics Letters*, 117, 794–98. 1

A Appendix

Table A.1: *List of Commodities*

#	Commodity (English)	Commodity (Japanese)	Wholesale Market Survey	Family Income and Expenditure Survey	Broad commodity category
Fruits:					
1	apple	りんご	✓	✓	fruit
2	cherry	おうとう	✓	✗	–
3	chestnut	くり	✓	✗	–
4	grape	ぶどう	✓	✓	fruit
5	Japanese apricot	うめ	✓	✗	–
6	Japanese plum	すもも	✓	✗	–
7	kaki	かき	✓	✓	fruit
8	kiwi fruit	キウイフルーツ	✓	✓	fruit
9	loquat	びわ	✓	✗	–
10	miscellaneous citrus fruits	みかん	✓	✓	fruit
11	muskmelon	メロン	✓	✓	fruit
12	nashi	なし	✓	✓	fruit
13	peach	もも	✓	✓	fruit
14	pear	西洋なし	✓	✗	–
15	Satsuma mandarin	蜜柑	✓	✗	–
16	strawberry	いちご	✓	✓	fruit
17	watermelon	すいか	✓	✓	fruit
Vegetables:					
1	asparagus	アスパラガス	✓	✗	–
2	bamboo shoot	たけのこ	✓	✓	vegetable (root)
3	bell pepper	ピーマン	✓	✓	vegetable (fruit)
4	bok choy	ちんげんさい	✓	✗	–
5	broad bean	そらまめ	✓	✗	–
6	broccoli	ブロッコリー	✓	✓	vegetable (leaf)
7	cabbage	キャベツ	✓	✓	vegetable (leaf)
8	carrot	にんじん	✓	✓	vegetable (root)
9	cauliflower	カリフラワー	✓	✗	–
10	celery	セルリー	✓	✗	–
11	cherry tomato	ミニトマト	✓	✗	–
12	cucumber	きゅうり	✓	✓	vegetable (fruit)
13	daikon radish	だいこん	✓	✓	vegetable (root)
14	eggplant	なす	✓	✓	vegetable (fruit)
15	enoki mushroom	えのきだけ	✓	✗	–
16	garland chrysanthemum	しゅんぎく	✓	✗	–
17	garlic	にんにく	✓	✗	–
18	garlic chives	にら	✓	✗	–
19	giant butterbur	ふき	✓	✗	–
20	ginger	しょうが	✓	✗	–

Continued on next page...

Table A.1: *List of Commodities (continued from previous page)*

#	Commodity (English)	Commodity (Japanese)	Wholesale Market Survey	Family Income and Expenditure Survey	Broad commodity category
21	great burdock	ごぼう	✓	✓	vegetable (root)
22	green beans	さやいんげん	✓	✗	–
23	green peas	実えんどう	✓	✗	–
24	green soybeans	えだまめ	✓	✗	–
25	lettuce	レタス	✓	✓	vegetable (leaf)
26	lotus root	れんこん	✓	✓	vegetable (root)
27	mizuna	みずな	✓	✗	–
28	mustard spinach	こまつな	✓	✗	–
29	nameko mushroom	なめこ	✓	✗	–
30	napa cabbage	はくさい	✓	✓	vegetable (leaf)
31	onion	たまねぎ	✓	✓	vegetable (root)
32	parsley	パセリ	✓	✗	–
33	potato	ばれいしょ	✓	✓	vegetable (root)
34	pumpkin	かぼちゃ	✓	✓	vegetable (fruit)
35	sugar peas	さやえんどう	✓	✗	–
36	shiitake mushroom	生しいたけ	✓	✓	vegetable (fruit)
37	shimeji mushroom	しめじ	✓	✗	–
38	small sweet green pepper	ししとうがらし	✓	✗	–
39	spikenard	うど	✓	✗	–
40	spinach	ほうれんそう	✓	✓	vegetable (leaf)
41	sweet corn	スイートコーン	✓	✗	–
42	sweet potato	かんしょ	✓	✓	vegetable (root)
43	taro	さといも	✓	✓	vegetable (root)
44	tomato	トマト	✓	✓	vegetable (fruit)
45	turnip	かぶ	✓	✗	–
46	welsh onion	ねぎ	✓	✓	vegetable (leaf)
47	wildparsley	みつば	✓	✗	–
48	yam	やまのいも	✓	✗	–

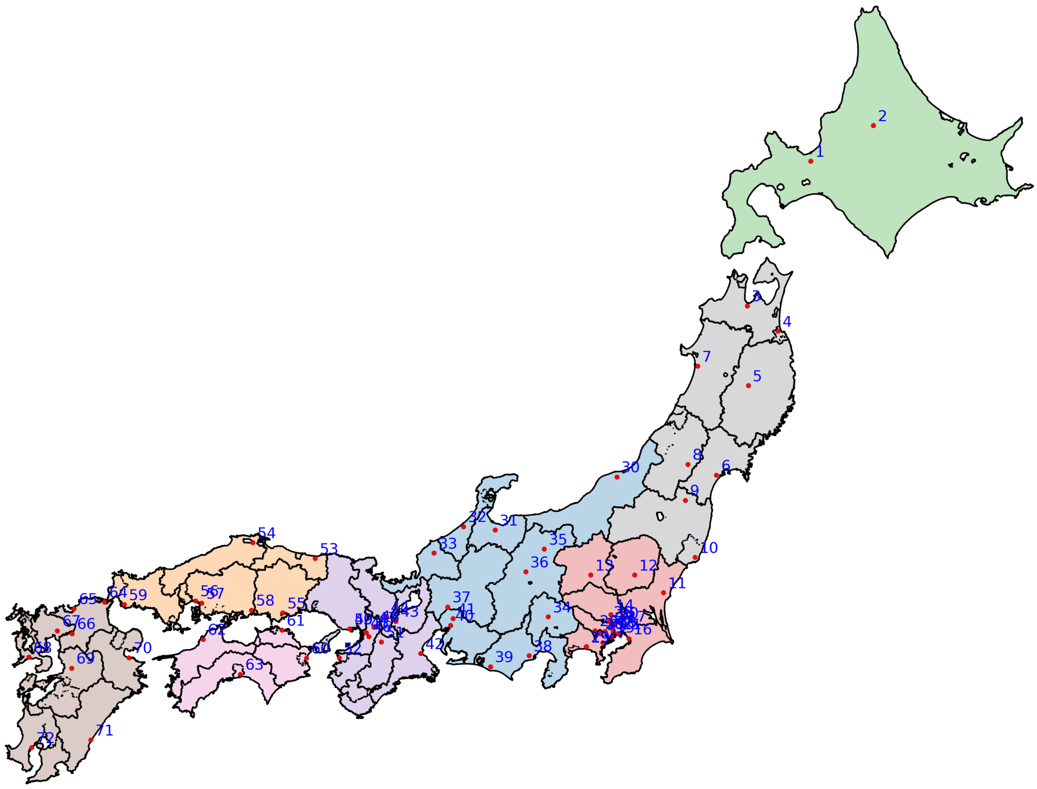
Notes: Table A.1 lists all 65 commodities covered by the Survey of Fresh Fruits and Vegetables Wholesale Market Survey and all 31 commodities covered by the Family Income and Expenditure Survey with their names (in English and Japanese). The 31 commodities from the Japanese Family Income and Expenditure Survey are classified into the following 4 broad goods categories: "fruit", "vegetable (root)", "vegetable (leaf)" and "vegetable (fruit)".

Table A.2: *List of Japanese Wholesale Markets for Fruits and vegetables*

#	Market (English)	Market (Japanese)	Prefecture	City	Municipality
1	Sapporo City Wholesale Market	札幌市中央卸売市場	Hokkaido	Sapporo	Sapporo
2	Asahikawa City Fruits and Vegetables Market	旭川市青果市場	Hokkaido	Asahikawa	Asahikawa
3	Aomori City Wholesale Market	青森市中央卸売市場	Aomori	Aomori	Aomori
4	Hachinohe City Wholesale Market	八戸市中央卸売市場	Aomori	Hachinohe	Hachinohe
5	Morioka City Wholesale Market	盛岡市中央卸売市場	Iwate	Morioka	Morioka
6	Sendai City Wholesale Market	仙台市中央卸売市場	Miyagi	Sendai	Sendai
7	Akita City Fruits and Vegetables Market	秋田市青果市場	Akita	Akita	Akita
8	Yamagata City Fruits and Vegetables Market	山形市青果市場	Yamagata	Yamagata	Yamagata
9	Fukushima City Fruits and Vegetables Market	福島市青果市場	Fukushima	Fukushima	Fukushima
10	Iwaki City Wholesale Market	いわき市中央卸売市場	Fukushima	Iwaki	Hitachi
11	Mito City Fruits and Vegetables Market	水戸市青果市場	Ibaraki	Mito	Mito
12	Utsunomiya City Wholesale Market	宇都宮市中央卸売市場	Tochigi	Utsunomiya	Utsunomiya
13	Maebashi City Fruits and Vegetables Market	前橋市青果市場	Gunma	Maebashi	Maebashi
14	Saitama City Wholesale Market	大宮総合食品卸売市場	Saitama	Saitama	Saitama
15	Tokyo City JA-Center	ＪＡ全農東京センター	Saitama	Toda	Toda
16	Chiba City Fruits and Vegetables Market	千葉市青果市場	Chiba	Chiba	Chiba
17	Matsudo City Fruits and Vegetables Market	松戸市青果市場	Chiba	Matsudo	Matsudo
18	Tokyo Prefecture Tsukiji Wholesale Market	東京都中央築地市場	Tokyo	Tokyo	Tokyo
19	Tokyo Prefecture Ota Wholesale Market	東京都中央大田市場	Tokyo	Tokyo	Tokyo
20	Tokyo Prefecture Kita-Adachi Wholesale Market	東京都中央北足立市場	Tokyo	Tokyo	Tokyo
21	Tokyo Prefecture Kasai Wholesale Market	東京都中央葛西市場	Tokyo	Tokyo	Tokyo
22	Tokyo Prefecture Toshima Wholesale Market	東京都中央豊島市場	Tokyo	Tokyo	Tokyo
23	Tokyo Prefecture Yodobashi Wholesale Market	東京都中央淀橋市場	Tokyo	Tokyo	Tokyo
24	Tokyo Prefecture Setagaya Wholesale Market	東京都中央世田谷市場	Tokyo	Tokyo	Tokyo
25	Tokyo Prefecture Itabashi Wholesale Market	東京都中央板橋市場	Tokyo	Tokyo	Tokyo
26	Tokyo Prefecture Tama Wholesale Market	東京都中央多摩市場	Tokyo	Tama	Tama
27	Yokohama City Central Wholesale Market	横浜市中央市場本場	Kanagawa	Yokohama	Yokohama
28	Kawasaki City Wholesale Market	川崎市中央卸売市場	Kanagawa	Kawasaki	Kawasaki
29	Kanagawa Prefecture JA-Center	ＪＡ全農神奈川センター	Kanagawa	Hiratsuka	Hiratsuka
30	Niigata City Wholesale Market	新潟市中央卸売市場	Niigata	Niigata	Niigata
31	Toyama City Fruit and Wholesale Market	富山市青果市場	Toyama	Toyama	Toyama
32	Kanazawa City Wholesale Market	金沢市中央卸売市場	Ishikawa	Kanazawa	Kanazawa
33	Fukui City Wholesale Market	福井市中央卸売市場	Fukui	Fukui	Fukui
34	Kofu City Fruits and Vegetables Market	甲府市青果市場	Yamanashi	Kofu	Kofu
35	Nagano City Fruits and Vegetables Market	長野市青果市場	Nagano	Nagano	Nagano
36	Matsumoto City Fruits and Vegetables Market	松本市青果市場	Nagano	Matsumoto	Matsumoto
37	Gifu City Wholesale Market	岐阜市中央卸売市場	Gifu	Gifu	Gifu
38	Shizuoka City Wholesale Market	静岡市中央卸売市場	Shizuoka	Shizuoka	Shizuoka
39	Hamamatsu City Wholesale Market	浜松市中央卸売市場	Shizuoka	Hamamatsu	Hamamatsu
40	Nagoya City Central Wholesale Market	名古屋市中央市場本場	Aichi	Nagoya	Nagoya
41	Nagoya City Northern Wholesale Market	名古屋市中央市場北部	Aichi	Nagoya	Nagoya
42	Mie Prefecture Fruits and Vegetables Market	三重県青果市場	Mie	Matsusaka	Matsusaka
43	Otsu City Fruits and Vegetables Market	大津市青果市場	Shiga	Otsu	Otsu
44	Kyoto City Wholesale Market	京都市中央卸売市場	Kyoto	Kyoto	Kyoto
45	Osaka City Central Wholesale Market	大阪市中央市場本場	Osaka	Osaka	Osaka
46	Osaka City Eastern Wholesale Market	大阪市中央市場東部	Osaka	Osaka	Osaka
47	Osaka Prefecture Wholesale Market	大阪府中央卸売市場	Osaka	Ibaraki	Ibaraki
48	Osaka City JA-Center	ＪＡ全農大阪センター	Osaka	Takatsuki	Takatsuki
49	Kobe City Central Wholesale Market	神戸市中央市場本場	Hyogo	Kobe	Kobe
50	Kobe City Eastern Wholesale Market	神戸市中央市場東部	Hyogo	Kobe	Kobe
51	Nara Prefecture Wholesale Market	奈良県中央卸売市場	Nara	Nara	Nara
52	Wakayama City Wholesale Market	和歌山市中央卸売市場	Wakayama	Wakayama	Wakayama
53	Tottori City Fruits and Vegetables Market	鳥取市青果市場	Tottori	Tottori	Tottori
54	Matsue City Fruits and Vegetables Market	松江市青果市場	Shimane	Matsue	Matsue
55	Okayama City Wholesale Market	岡山市中央卸売市場	Okayama	Okayama	Okayama
56	Hiroshima City Central Wholesale Market	広島市中央市場中央	Hiroshima	Hiroshima	Hiroshima
57	Hiroshima City Eastern Wholesale Market	広島市中央市場東部	Hiroshima	Hiroshima	Hiroshima
58	Fukuyama City Fruits and Vegetables Market	福山市青果市場	Hiroshima	Fukuyama	Fukuyama
59	Ube City Wholesale Market	宇部市中央卸売市場	Yamaguchi	Ube	Ube
60	Tokushima City Wholesale Market	徳島市中央卸売市場	Tokushima	Tokushima	Tokushima
61	Takamatsu City Wholesale Market	高松市中央卸売市場	Kagawa	Takamatsu	Takamatsu
62	Matsuyama City Wholesale Market	松山市中央卸売市場	Ehime	Matsuyama	Matsuyama
63	Kochi City Wholesale Market	高知市中央卸売市場	Kochi	Kochi	Kochi
64	Kitakyushu City Wholesale Market	北九州市中央卸売市場	Fukuoka	Kitakyushu	Kitakyushu
65	Fukuoka City Wholesale Market	福岡市中央卸売市場	Fukuoka	Fukuoka	Fukuoka
66	Kurume City Wholesale Market	久留米市中央卸売市場	Fukuoka	Kurume	Kurume
67	Saga City Fruits and Vegetables Market	佐賀市青果市場	Saga	Saga	Saga
68	Nagasaki City Wholesale Market	長崎市中央卸売市場	Nagasaki	Nagasaki	Nagasaki
69	Kumamoto City Fruits and Vegetables Market	熊本青果市場	Kumamoto	Kumamoto	Kumamoto
70	Oita City Fruits and Vegetables Market	大分市青果市場	Oita	Oita	Oita
71	Miyazaki City Wholesale Market	宮崎市中央卸売市場	Miyazaki	Miyazaki	Miyazaki
72	Kagoshima City Wholesale Market	鹿児島市中央卸売市場	Kagoshima	Kagoshima	Kagoshima

Notes: Table A.2 list 72 wholesale markets for fruits and vegetables with their names (in English and Japanese) and their location (i.e. prefecture and city). To match the Fresh Fruits and Vegetables Wholesale Market Survey to the Family Income and Expenditure Survey each wholesale market is associated with the geographically closest municipality from the Family Income and Expenditure Survey.

Figure A.1: *Wholesale markets for fruits and vegetables in Japan*



Notes: Figure A.1 plots the geographic distribution of Japan's 72 wholesale markets for fruits and vegetables, which are numbered as in Table A.2.

Table A.3: *List of survey items included in the LAJDB*

Item #	Map #	Question (English)	Question (Japanese)
1	229, 230	praying mantis	かまきり（蟷螂）
2	233	spider	くも（蜘蛛）
3	235	thread (of a spider's web)	くものいと（蜘蛛の糸）
4	234	spider's web	くものす（蜘蛛の巣）
5	235, 236, 237	snail	かたつむり（蝸牛）
6	239	slug	なめくじ（蛞蝓）
7	221, 222, 223	tadpole	おたまじゃくし（蝌蚪）
8	218	frog	かえる（蛙）
12	224	black lizard	とかげ（蜥蜴）
21	91, 92	to tell a lie	うそ（嘘言）をつく
23	56	to make	つくる（作る）
31	101	head	あたま（頭）
32	102	wirl of hair at the head	つむじ（旋毛）
34	110	eye	め（目）
36	112	sty	ものもらい（麦粒腫）
37	113	nose	はな（鼻）
38	268	pleasant smell	におい（芳香）
39	269	bad smell	におい（悪臭）
42	85, 86	to smell	におい（匂）をかぐ（嗅ぐ）
44	115	mouth	くち（口）
47	116	lips	くちびる（唇）
48	117	tongue	した（舌）
51	38	weak, thin (taste)	〈塩味が〉うすい
52	37	sweet (taste)	あまい（甘い）
56	107	cheek	ほほ（頬）
57	106	face	かお（顔）
59	80	to become black and blue (after a blow)	あざ（痣）になる
60	133	small mole	ほくろ（黒子） 小さいもの
61	134	large mole	ほくろ（黒子） 大きいもの
63	121	thumb	おやゆび（親指）
64	122	index finger	ひとさしゆび（人差し指）
65	123	middle finger	なかゆび（中指）
66	124	ring finger	くすりゆび（薬指）
67	125	little finger	こゆび（小指）
68	127	chilblains	しもやけ（凍傷）
69	129	heel	かかと（踵）
72	51	to sit on one's heels (to kneel on the floor)	すわる（坐る）
73	130	solar plexus	みずおち（鳩尾）
74	131	body dirt	あか（垢）
75	105	dandruff	ふけ（雲脂）
76	217	scale	うろこ（鱗）
79	245	mushroom	きのこ（茸・蕈）
80	136	man	おとこ（男）
81	137	woman	おんな（女）
83	144	stilts	たけうま（竹馬）
89	148	hide-and-seek	かくれんぼ（隠れん坊）
90	151	money	おかね（貨幣）
91	152	(to give back) the change	おつり（釣銭）
92	296	to count (money)	かぞえる（お金を数える）
93	69	to count	かぞえる（数える）
95	73	to give (toward the hearer)	やる（遣る）
102	280	today	きょう（今日）
103	278	yesterday	きのう（昨日）
104	276	the day before yesterday	おととい（一昨日）
105	275	two days before yesterday	さきおととい（一昨昨日）
108	282	tomorrow	あした（明日）
109	284	the day after tomorrow	あさって（明後日）
110	285	two days after tomorrow	しあさって（明明後日）

Continued on next page...

Table A.3: *List of survey items included in the LAJDB (continued from previous page)*

Question #	Map #	Question (English)	Question (Japanese)
111	286	three days after tomorrow	やのあさって（明明明後日）
112	281	tonight	こんばん（今晩）
113	283	tomorrow night	あしたのばん（明晩）
114	251	sun	たいよう（太陽）
116	252	moon	つき（月）
117	253	rain	あめ（雨）
118	254	rain (during the raining season)	つゆ（梅雨）
119	255	shower (rain)	ゆうだち（夕立雨）
122	258	lightning	いなずま（稲妻・電光）
124	259	rainbow	にじ（虹）
125	260	snow	ゆき（雪）
127	96	to be frozen (water)	こおる（水が凍る）
129	262	icicle	つらら（氷柱）
131	272	grit in the eye	ごみ（目にはいるもの　塵）
132	273	sweepings	ごみ（掃除の対象　塵芥）
134	274	river waste	ごみ（川のごみ　塵芥）
135	263	earthquake	じしん（地震）
148	57	to boil rice	たく（炊く）
149	58	to boil vegetables	にる（煮る）
153	267	steam (from cooked rice)	ゆげ（蒸気　飯の場合）
155	162	earthenware mortar	すりばち（擂鉢）
156	163	wooden pestle	すりこぎ（擂粉木）
157	161	china, porcelain	せともの（陶磁器）
164	158	cotton wool	わた（綿）
165	159	floss-silk, silk-wadding	まわた（真綿）
166	153	thread	いと（糸）
167	155	silk thread	きぬいと（絹糸）
169	157	weaving thread, strand	はたいと（機糸）
173	167	rice	こめ（米）
174	168	nonglutinous rice	うるち（粳米）
176	170	rice for the farmer's private consumption	はんまい（飯米）
179	172	rice-bran	ぬか（糠）
182	187	ridge between rice-fields	あぜ（畦畔）
184	189	scarecrow (noise or light making device)	とりおどし（鳥威）
185	190	scarecrow (human figure)	かかし（案山子）
186	174, 175	potato	じゃがいも（馬鈴薯）
187	177, 178	taro	さといも（里芋）
188	176	sweet potato	さつまいも（甘藷）
190	182	indian corn, maize	とうもろこし（玉蜀黍）
191	180	pumpkin	かぼちゃ（南瓜）
192	240	violet (flower)	すみれ（菫）
194	244	horsetail (cone)	つくし（土筆）
195	243	horsetail	すぎな（杉菜・問荊）
197	297	water well	まつかさ（松毬）
200	250	thorn	とげ（刺・棘）　いばらやさんしょうなどのとげ
213	201	horse	うま（馬）
214	202	horse, stallion	おうま（牡馬）
215	203	mare	めうま（牝馬）
216	204	foal	こうま（子馬）
217	205	mane (of the horse)	たてがみ（鬣）
218	206	cattle	うし（牛）
219	207	bull	おうし（牡牛）
220	208	cow	めうし（牝牛）
221	209	calf	こうし（子牛）
222	210	moo (the lowing of cattle)	もうもう（牛の鳴き声）
223	211	mole (animal)	もぐら（土竜・〓鼠）
224	212	owl	ふくろう（梟）
228	214	sparrow	すずめ（雀）

Continued on next page...

Table A.3: *List of survey items included in the LAJDB (continued from previous page)*

Question #	Map #	Question (English)	Question (Japanese)
229	300	chirp (the cry of the sparrow)	ちゅんちゅん（雀の鳴き声）
231	104	to grow bald	はげる（禿げる）
233	128	malleolus	くるぶし（踝）
235	62	to trow away	すてる（捨てる）
236	77	to be surprised	びっくりする（驚く）
237	42	fearful, dreadful	おそろしい（恐ろしい）
238	287	the seventh day (of the month)	なのか（七日）
239	288	the ninth day (of the month)	ここのか（九日）
240	139	great-grandchild	ひまご（曾孫）
241	140	great great grandchild	やしゃご（玄孫）
244	191	house	いえ（家屋）
248	192	sliding door between rooms	ふすま（襖障子）
250	47	beautiful, lovely	〈虹が〉きれいだ
252	183	red pepper	とうがらし（蕃椒）
253	291	tasty	おいしい（美味しい）
258	12	HI in HIGE (mustache)	ヒゲ（鬚）の HI-の音
261	64	to carry (a baby on one's back)	おんぶする（幼児を負う）
264	66	to carry (a log on one's shoulder)	かつぐ（材木を担ぐ）
265	67	to carry (on both ends of a pole)	かつぐ（天秤棒を担ぐ）
266	68	to carry (suspended on a pole hanging between two people)	かつぐ（二人で担ぐ）
268	53	to be, to exist	いる（居る）
270	45, 46	(it is) fine Weather), (it) is (fine weather)	いい〈天気だ〉
280	1	G in KAGAMI (mirror)	カガミ（鏡）の-G-の音
282	181	eggplant	なす（茄子）
284	231	dragonfly	とんぼ（蜻蛉）

Notes: Table A.3 lists 141 questionnaire items with the corresponding map numbers from the Linguistic Atlas of Japan (LAJ) that are contained in the Linguistic Atlas of Japan Database (LAJDB).

Table A.4: *Integration of Japanese wholesale markets for fruits and vegetables*

Dependent variable: Relative price volatility for commodity c b/w the markets i and j ($REL\_PRICE\_VOL_{ijc}$)			
Model:	**OLS-FE**		
Specification:	**(1)**	**(2)**	**(3)**
$\ln DIST_{ij}$	0.0148***	0.0090***	0.0096***
	(.0014)	(.0014)	(.0016)
$JOINT\_SUPPLIERS_{ijc}$		−0.0642***	−0.0633***
		(.0133)	(.0135)
$PREF\_BORD_{ij}$			−0.0044*
			(.0024)
$REGION\_BORD_{ij}$			−0.0023*
			(.0012)
$ISLAND\_BORD_{ij}$			0.0054**
			(.0025)
Fixed effects:			
Market-specific fixed effects:	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓
Summary statistics:			
# of observations:	158,472	158,472	158,472
R^2:	.8182	.8208	.8209

Notes: Robust standard errors clustered at the commodity level; significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.5: *AIDS estimation results for vegetables (root)*

Dependent variable: Budget share s_{hct} of household h spend on commodity c at time $t = 2019$.									
Model:	**OLS-FE**								
Specification:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Commodity:	daikon radish	carrot	great burdock	bamboo shoot	lotus root	sweet potato	potato	taro	onion
ln total vegetable (root) expenditure	-0.0001^{***}	-0.0001^{***}	0.0000	0.0001^{***}	0.0001^{***}	0.0001^{***}	0.0000^{***}	0.0001^{***}	-0.0001^{***}
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
ln price daikon radish	-0.0013	-0.0062^{**}	-0.0049^{**}	0.0106^{***}	-0.0017	-0.0022	0.0040	-0.0035	0.0052
	(.0032)	(.0031)	(.0024)	(.0030)	(.0023)	(.0019)	(.0044)	(.0028)	(.0054)
ln price carrot	0.0124^{**}	-0.0014	-0.0024	0.0016	0.0164^{***}	0.0026	-0.0180^{**}	0.0146^{***}	-0.0258^{***}
	(.0053)	(.0040)	(.0030)	(.0042)	(.0034)	(.0035)	(.0074)	(.0054)	(.0085)
ln price great burdock	-0.0204^{***}	-0.0026	-0.0037	-0.0002	-0.0060^{**}	-0.0047	0.0254^{***}	-0.0162^{***}	0.0285^{***}
	(.0061)	(.0027)	(.0028)	(.0030)	(.0030)	(.0029)	(.0057)	(.0043)	(.0082)
ln price bamboo shoot	-0.0025	-0.0024^{**}	0.0013	-0.0021	0.0028^{**}	-0.0005	0.0016	0.0032^{**}	-0.0015
	(.0019)	(.0011)	(.0010)	(.0014)	(.0012)	(.0010)	(.0016)	(.0015)	(.0025)
ln price lotus root	-0.0131^{***}	0.0186^{***}	0.0025	0.0185^{***}	-0.0269^{***}	-0.0176^{***}	0.0367^{***}	-0.0181^{***}	-0.0005
	(.0048)	(.0045)	(.0039)	(.0038)	(.0038)	(.0032)	(.0063)	(.0041)	(.0085)
ln price sweet potato	0.0243^{**}	0.0355^{***}	-0.0039	-0.0265^{***}	-0.0143^{***}	0.0129^{***}	0.0159	-0.0096	-0.0343^{**}
	(.0123)	(.0084)	(.0061)	(.0089)	(.0048)	(.0045)	(.0116)	(.0106)	(.0147)
ln price potato	-0.0213^{***}	0.0020	-0.0092^{**}	0.0203^{***}	-0.0239^{***}	-0.0100^{**}	0.0236^{***}	-0.0280^{***}	0.0465^{***}
	(.0077)	(.0048)	(.0041)	(.0052)	(.0041)	(.0042)	(.0065)	(.0053)	(.0102)
ln price taro	0.0065	0.0163^{***}	0.0091^{**}	-0.0078^{*}	-0.0036	-0.0045	-0.0038	-0.0017	-0.0104
	(.0050)	(.0043)	(.0040)	(.0043)	(.0040)	(.0030)	(.0067)	(.0044)	(.0072)
ln price onion	-0.0169^{***}	-0.0090^{**}	-0.0001	0.0043	-0.0035	0.0091^{***}	0.0089	-0.0091	0.0164^{*}
	(.0058)	(.0037)	(.0031)	(.0039)	(.0044)	(.0029)	(.0058)	(.0063)	(.0086)

Notes: Table A.5 contains the results of the AIDS OLS estimation of the demand for 9 vegetables (root) in the 2019 Japanese Family Income and Expenditure Survey based on Eq. (11). The dependent variable is the budget share s_{hct} that household h spends on commodity c in survey year $t = 2019$. ln total vegetable (root) expenditure is the log of total expenditure on root vegetables divided by a municipality-level Stone price index. ln price by commodity category c is the median price of commodity c across all wholesale markets in municipality m. All regressions include a vector of household characteristics (number of household members, number of household members below age 18 and above age 65, age and employment status of the household head, occupation-specific and industry-specific fixed effect), as well as a municipality-specific fixed effect which deliver our regional taste estimates. The estimation is based on 427,290 observations and has an $R^2 = 0.2442$. Standard errors are clustered at the municipality level and reported in parenthesis with *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.6: *AIDS estimation results for vegetables (leaf)*

Dependent variable: Budget share s_{hct} of household h spend on commodity c at time $t = 2019$.						
Model:	**OLS-FE**					
Specification:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Commodity:	napa cabbage	cabbage	spinach	welsh onion	broccoli	lettuce
ln total vegetable (leaf) expenditure	-0.0000^{***}	-0.0002^{***}	0.0001^{***}	-0.0000^{***}	0.0002^{***}	0.0000
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
ln price napa cabbage	-0.0360^{***}	0.0380^{***}	-0.0233^{**}	-0.0212^{***}	-0.0032	0.0457^{***}
	(.0117)	(.0122)	(.0110)	(.0078)	(.0059)	(.0148)
ln price cabbage	-0.0442^{***}	0.0480^{***}	0.0076	-0.0385^{***}	-0.0143^{*}	0.0415^{***}
	(.0088)	(.0100)	(.0064)	(.0094)	(.0077)	(.0092)
ln price spinach	0.0558^{***}	-0.0142	-0.0738^{***}	0.0442^{***}	-0.0098	-0.0022
	(.0112)	(.0130)	(.0104)	(.0112)	(.0090)	(.0108)
ln price welsh onion	0.0358^{***}	-0.0105	-0.0105	0.0028	-0.0033	-0.0142^{**}
	(.0082)	(.0083)	(.0074)	(.0061)	(.0055)	(.0062)
ln price broccoli	-0.0370^{***}	0.0568^{***}	-0.0096	-0.0068	-0.0429^{***}	0.0395^{***}
	(.0096)	(.0098)	(.0080)	(.0082)	(.0094)	(.0126)
ln price lettuce	0.0228^{***}	-0.0057	0.0072	0.0101^{*}	-0.0076	-0.0268^{***}
	(.0082)	(.0073)	(.0050)	(.0061)	(.0055)	(.0075)

Notes: Table A.6 contains the results of the AIDS OLS estimation of the demand for 6 vegetables (leaf) in the 2019 Japanese Family Income and Expenditure Survey based on Eq. (11). The dependent variable is the budget share s_{hct} that household h spends on commodity c in survey year $t = 2019$. ln total vegetable (leaf) expenditure is the log of total expenditure on leaf vegetables divided by a municipality-level Stone price index. ln price by commodity category c is the median price of commodity c across all wholesale markets in municipality m. All regressions include a vector of household characteristics (number of household members, number of household members below age 18 and above age 65, age and employment status of the household head, occupation-specific and industry-specific fixed effect), as well as a municipality-specific fixed effect which deliver our regional taste estimates. The estimation is based on 333,132 observations and has an $R^2 = 0.0917$. Standard errors are clustered at the municipality level and reported in parenthesis with *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.7: *AIDS estimation results for vegetables (misc.)*

Dependent variable: Budget share s_{hct} of household h spend on commodity c at time $t = 2019$.						
Model:	**OLS-FE**					
Specification:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Commodity:	cucumber	pumkin	eggplant	tomato	bell pepper	shiitake mushroom
ln total vegetable (misc.) expenditure	-0.0001^{***}	-0.0000^{***}	0.0000^{*}	0.0003^{***}	-0.0001^{***}	-0.0001^{***}
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
ln price cucumber	-0.0469^{***}	0.0157^{**}	-0.0321^{***}	-0.0145^{*}	-0.0061	0.0838^{***}
	(.0076)	(.0063)	(.0085)	(.0076)	(.0059)	(.0102)
ln price pumkin	-0.0091^{**}	-0.0108^{***}	-0.0307^{***}	0.0289^{***}	0.0005	0.0211^{***}
	(.0037)	(.0031)	(.0039)	(.0042)	(.0024)	(.0048)
ln price eggplant	-0.0012	0.0100	-0.0436^{***}	0.0066	-0.0096^{*}	0.0378^{***}
	(.0083)	(.0061)	(.0094)	(.0115)	(.0051)	(.0135)
ln price tomato	0.0161^{***}	0.0043	-0.0149^{***}	-0.0339^{***}	0.0012	0.0272^{***}
	(.0048)	(.0042)	(.0046)	(.0087)	(.0043)	(.0054)
ln price bell pepper	0.0036	-0.0029	-0.0080	0.0125	-0.0020	-0.0032
	(.0082)	(.0056)	(.0079)	(.0078)	(.0047)	(.0079)
ln price shiitake mushroom	0.0178^{*}	-0.0085	-0.0509^{***}	0.0081	-0.0077	0.0412^{***}
	(.0094)	(.0072)	(.0110)	(.0113)	(.0070)	(.0115)

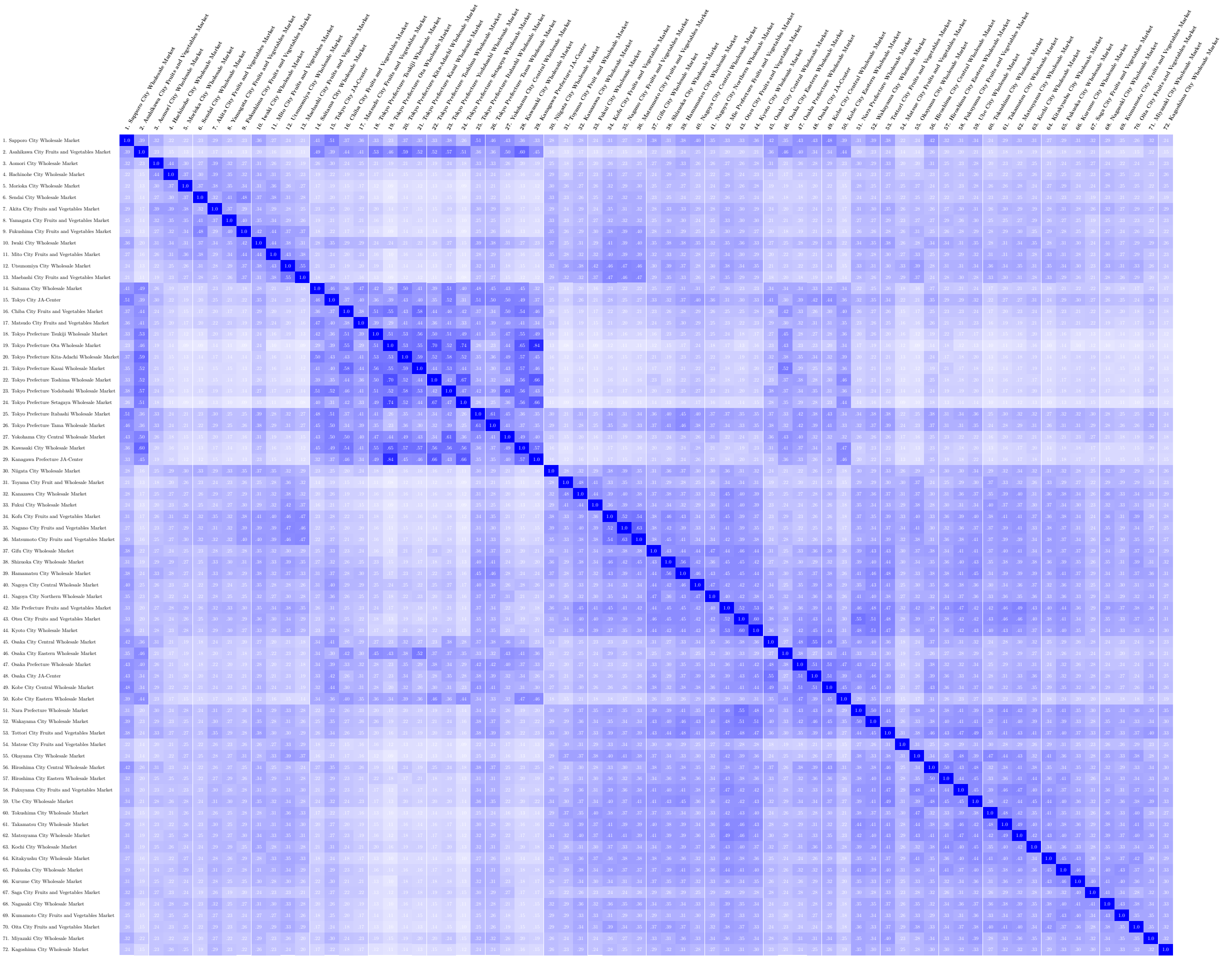
Notes: Table A.7 contains the results of the AIDS OLS estimation of the demand for 6 vegetables (misc.) in the 2019 Japanese Family Income and Expenditure Survey based on Eq. (11). The dependent variable is the budget share s_{hct} that household h spends on commodity c in survey year $t = 2019$. ln total vegetable (misc.) expenditure is the log of total expenditure on miscellaneous vegetables divided by a municipality-level Stone price index. ln price by commodity category c is the median price of commodity c across all wholesale markets in municipality m. All regressions include a vector of household characteristics (number of household members, number of household members below age 18 and above age 65, age and employment status of the household head, occupation-specific and industry-specific fixed effect), as well as a municipality-specific fixed effect which deliver our regional taste estimates. The estimation is based on $327,588$ observations and has an $R^2 = 0.2199$. Standard errors are clustered at the municipality level and reported in parenthesis with *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.8: *AIDS estimation results for fruits*

Dependent variable: Budget share s_{hct} of household h spend on commodity c at time $t = 2019$.										
Model:	**OLS-FE**									
Specification:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Commodity:	misc. citrus fruits	apple	pear	kaki	peach	grape	strawberry	muskmelon	watermelon	kiwi fruit
ln total fruit expenditure	0.0000^{***}	-0.0000^{***}	0.0000	-0.0000^{***}	0.0000^{***}	0.0000^{***}	-0.0000^{***}	0.0000^{**}	-0.0000^{***}	-0.0000^{***}
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
ln price misc. citrus fruits	0.0076^{*}	-0.0604^{***}	-0.0128^{**}	-0.0465^{***}	0.0137^{***}	0.0076^{*}	0.0302^{**}	0.0071^{*}	0.0351^{***}	0.0185^{***}
	(.0045)	(.0085)	(.0061)	(.0069)	(.0044)	(.0045)	(.0133)	(.0042)	(.0058)	(.0057)
ln price apple	-0.0657^{***}	-0.0051	-0.1371^{***}	0.0388^{**}	0.0573^{***}	-0.0657^{***}	-0.1792^{***}	0.0890^{***}	0.1595^{***}	0.1081^{***}
	(.0127)	(.0230)	(.0229)	(.0195)	(.0126)	(.0127)	(.0331)	(.0147)	(.0277)	(.0178)
ln price pear	-0.0013	-0.0167^{*}	-0.0201^{***}	-0.0418^{***}	-0.0024	-0.0013	0.0154	0.0129^{***}	0.0233^{***}	0.0321^{***}
	(.0048)	(.0096)	(.0068)	(.0075)	(.0042)	(.0048)	(.0107)	(.0041)	(.0055)	(.0043)
ln price kaki	0.0025	-0.0487^{***}	0.0187^{***}	-0.0595^{***}	0.0401^{***}	0.0025	0.0212^{*}	-0.0014	0.0377^{***}	-0.0132^{**}
	(.0045)	(.0120)	(.0057)	(.0113)	(.0091)	(.0045)	(.0123)	(.0038)	(.0081)	(.0062)
ln price peach	-0.0189^{***}	0.0177^{*}	-0.0284^{***}	-0.0205^{**}	-0.0091^{**}	-0.0189^{***}	0.1106^{***}	-0.0009	-0.0206^{***}	-0.0111^{*}
	(.0043)	(.0099)	(.0061)	(.0085)	(.0041)	(.0043)	(.0134)	(.0040)	(.0051)	(.0062)
ln price grape	-0.0126^{***}	0.0169^{***}	-0.0086^{**}	0.0082^{**}	-0.0131^{***}	-0.0126^{***}	0.0246^{***}	0.0053^{*}	-0.0129^{***}	0.0047
	(.0029)	(.0064)	(.0035)	(.0041)	(.0022)	(.0029)	(.0084)	(.0029)	(.0038)	(.0040)
ln price strawberry	0.0115	-0.0047	0.0194	0.0190^{**}	0.0175^{***}	0.0115	-0.0336^{**}	-0.0162^{**}	-0.0055	-0.0189^{***}
	(.0091)	(.0120)	(.0133)	(.0094)	(.0047)	(.0091)	(.0138)	(.0065)	(.0077)	(.0064)
ln price muskmelon	-0.0405^{***}	0.0739^{***}	-0.0358^{***}	-0.0402^{***}	0.0039	-0.0405^{***}	0.2001^{***}	-0.0265^{***}	-0.0330^{***}	-0.0614^{***}
	(.0083)	(.0182)	(.0123)	(.0133)	(.0070)	(.0083)	(.0227)	(.0071)	(.0093)	(.0097)
ln price watermelon	-0.0381^{***}	0.0857^{***}	-0.0704^{***}	0.0414^{***}	-0.0298^{***}	-0.0381^{***}	0.0585^{***}	0.0024	-0.0226^{***}	0.0109
	(.0056)	(.0136)	(.0109)	(.0101)	(.0059)	(.0056)	(.0159)	(.0046)	(.0070)	(.0069)
ln price kiwi fruit	0.0033	-0.0045	0.0022	0.0154^{*}	-0.0014	0.0033	0.0117	-0.0087	-0.0169^{***}	-0.0045
	(.0057)	(.0115)	(.0109)	(.0092)	(.0038)	(.0057)	(.0125)	(.0055)	(.0043)	(.0045)

Notes: Table A.8 contains the results of the AIDS OLS estimation of the demand for 10 fruits in the 2019 Japanese Family Income and Expenditure Survey based on Eq. (11). The dependent variable is the budget share s_{hct} that household h spends on commodity c in survey year $t = 2019$. ln total fruit expenditure is the log of total expenditure on fruits divided by a municipality-level Stone price index. ln price by commodity category c is the median price of commodity c across all wholesale markets in municipality m. All regressions include a vector of household characteristics (number of household members, number of household members below age 18 and above age 65, age and employment status of the household head, occupation-specific and industry-specific fixed effect), as well as a municipality-specific fixed effect which deliver our regional taste estimates. The estimation is based on 427,290 observations and has an $R^2 = 0.2442$. Standard errors are clustered at the municipality level and reported in parenthesis with *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

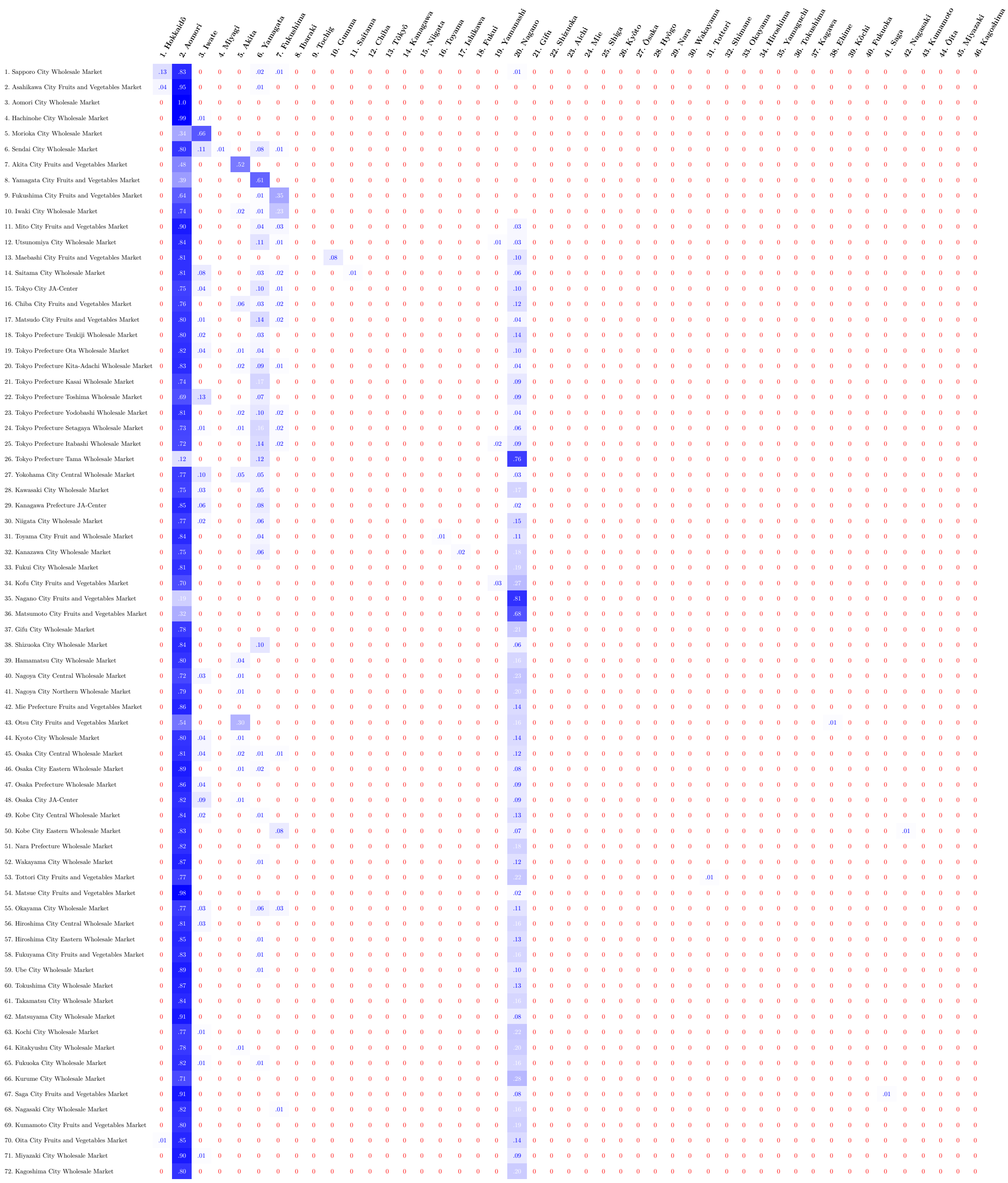
Table A.9: *The historical dialect similarity matrix*



Notes: Table A.9 plots the historical dialect similarity index $HIST\_DIALECT\_SIM_{ij} \in [0, 1]$ from Eq.(16) for each pair of the 72 Japanese wholesale markets for fruits and vegetables from Table A.2. By construction the matrix in Table A.9 is symmetric with respect to its main diagonal. Darker (lighter) shades of blue are associated with larger (smaller) index values. The 72 wholesale markets for fruits and vegetables are ordered from North-West (1. Sapporo City Wholesale Market) to South-East (72. Kagoshima City Wholesale Market).

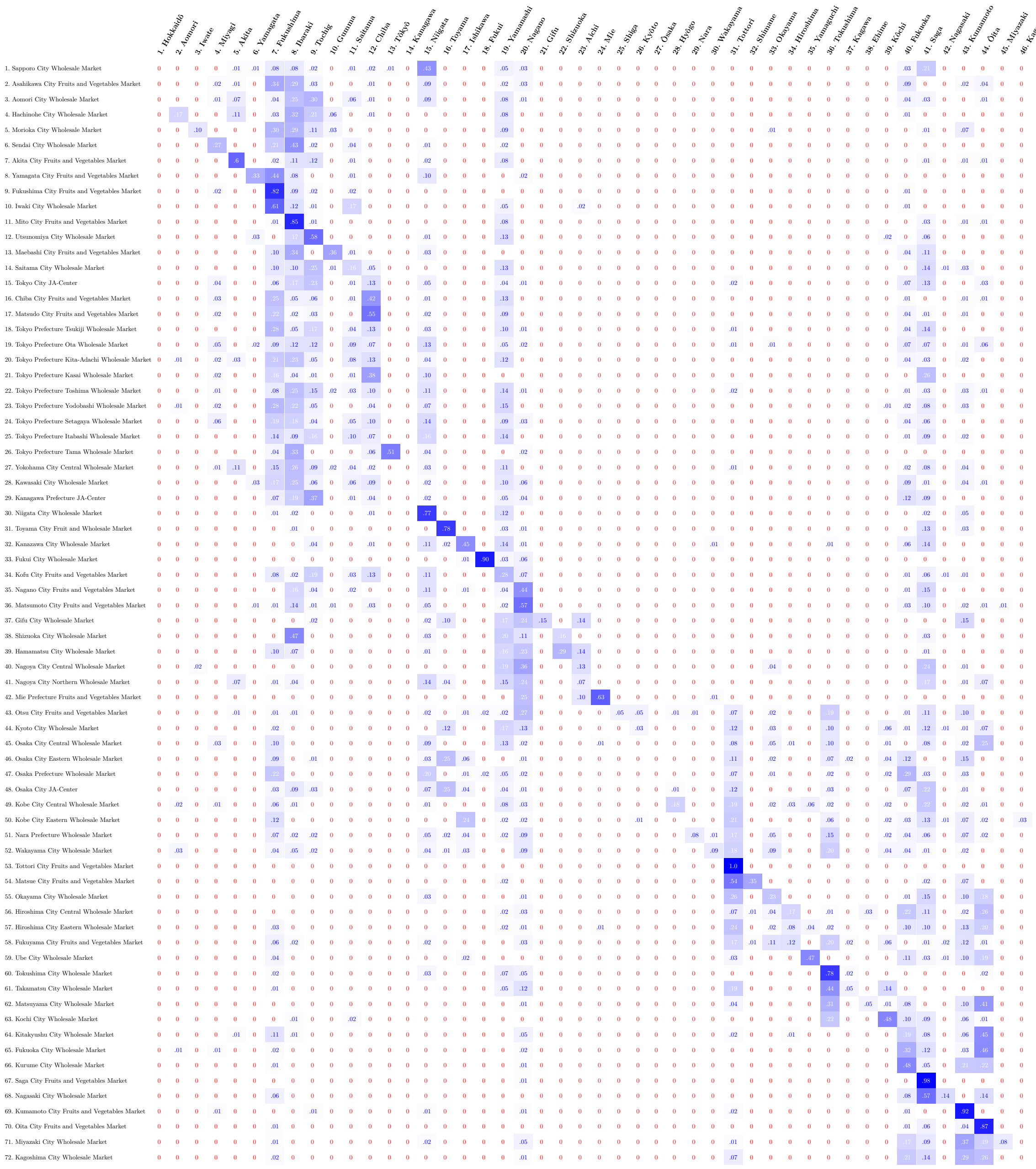
Figure A.2: *"Isaac Newton, discovering gravity while watching the fall"* from Nakagane (1873)

Figure A.3: *Average supply shares of apple producing prefectures (2010-2021)*



Notes: Figure A.3 plots for each of 72 wholesale markets for fruits and vegetables the average apple supply shares of the 46 producing prefecture from 2010 to 2021. Darker (lighter) shades of blue are associated with larger (smaller) shares. The 72 wholesale markets for fruits and vegetables are ordered from North-West (1. Sapporo City Wholesale Market) to South-East (72. Kagoshima City Wholesale Market). The 46 producing prefectures are also ordered from North-West (1. Hokkaidô) to South-East (46. Kagoshima).

Figure A.4: *Average supply shares of nashi producing prefectures (2010-2021)*



Notes: Figure A.4 plots for each of 72 wholesale markets for fruits and vegetables the average apple supply shares of the 46 producing prefecture from 2010 to 2021. Darker (lighter) shades of blue are associated with larger (smaller) shares. The 72 wholesale markets for fruits and vegetables are ordered from North-West (1. Sapporo City Wholesale Market) to South-East (72. Kagoshima City Wholesale Market). The 46 producing prefectures are also ordered from North-West (1. Hokkaidô) to South-East (46. Kagoshima).

Table A.10: *Accounting for unobserved heterogeneity (OLS)*

Dependent variable: Relative price volatility for commodity c b/w the markets i and j ($REL\_PRICE\_VOL_{ijc}$)				
Model:	**OLS-FE**			
Specification:	**(1)**	**(2)**	**(3)**	**(4)**
$TASTE\_DIST_{ijc}$	0.0741	0.0710	0.0679	0.0624
	(.0650)	(.0650)	(.0645)	(.0637)
$\ln DIST_{ij}$	0.0094***	0.0090***	0.0061***	0.0038***
	(.0016)	(.0016)	(.0010)	(.0013)
$JOINT\_SUPPLIERS_{ijc}$	−0.0631***	−0.0631***	−0.0598***	−0.0605***
	(.0135)	(.0136)	(.0145)	(.0154)
$PREF\_BORD_{ij}$	−0.0049**	−0.0051**	−0.0002	
	(.0023)	(.0023)	(.0015)	
$REGION\_BORD_{ij}$	−0.0023*	−0.0018		
	(.0012)	(.0012)		
$ISLAND\_BORD_{ij}$	0.0053**			
	(.0025)			
Fixed effects:				
Market-specific fixed effects:	✓	✓	✓	✓
Commodity-specific fixed effects:	✓	✓	✓	✓
Island-pair-specific fixed effects:	✗	✓	✓	✓
Region-pair-specific fixed effects:	✗	✗	✓	✓
Prefecture-pair-specific fixed effects:	✗	✗	✗	✓
Summary statistics:				
# of observations:	158,472	158,472	158,472	158,472
R^2:	.8209	.8210	.8212	.8221

Notes: Robust standard errors clustered at the commodity level; significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.11: *Zero-first-stage-test: hamburgers*

Dependent variable: Relative price volatility of hamburgers b/w the markets i and j ($REL\_PRICE\_VOL_{ij}$)				
Model:	**OLS-FE**			
Specification:	**(1)**	**(2)**	**(3)**	**(4)**
$HIST\_DIALECT\_SIM_{ij}$	0.0028	0.0026	0.0028	−0.0130*
	(.0045)	(.0046)	(.0048)	(.0073)
$\ln DIST_{ij}$	0.0011	0.0011	−0.0024**	−0.0049
	(.0007)	(.0007)	(.0011)	(.0032)
$PREF\_BORD_{ij}$	0.0075*	0.0095**	0.0127***	
	(.0041)	(.0041)	(.0039)	
$REGION\_BORD_{ij}$	−0.0007	−0.0004		
	(.0015)	(.0015)		
$ISLAND\_BORD_{ij}$	0.0014			
	(.0016)			
Fixed effects:				
Market-specific fixed effects:	✓	✓	✓	✓
Island-pair-specific fixed effects:	✗	✓	✓	✓
Region-pair-specific fixed effects:	✗	✗	✓	✓
Prefecture-pair-specific fixed effects:	✗	✗	✗	✓
Summary statistics:				
# of observations:	4,882	4,882	4,882	4,882
R^2:	.9536	.9538	.9555	.9804

Notes: Robust standard errors; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

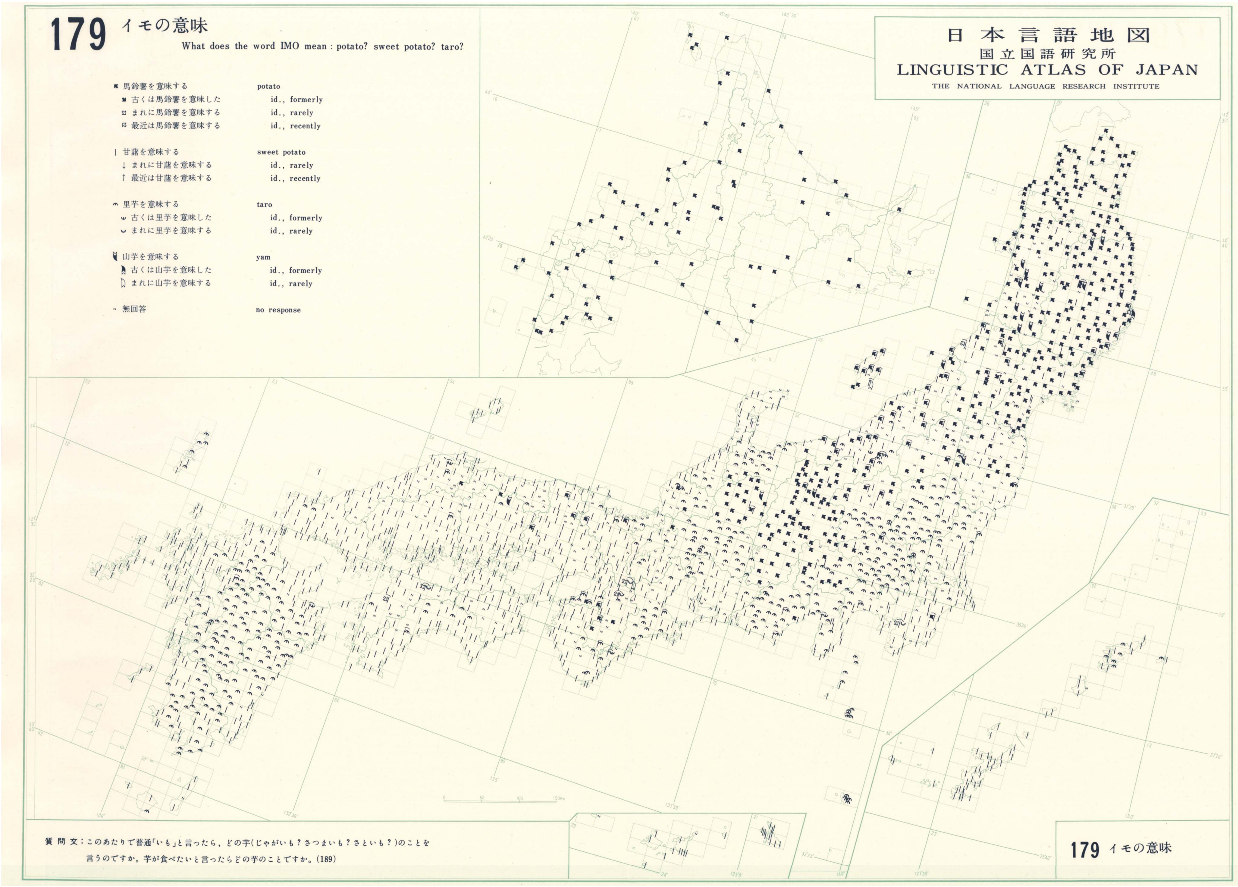


Figure A.5: *LAJ Map #179: Meaning of potato*