Subject to Availability

Option Production under Complementary Inputs

Online Appendix

Chae-Yeon Xon

Yonsei University

September 2026

Abstract

This appendix contains proofs, data construction, and additional results for the paper. The notation and target definitions follow the main text.

JEL Codes: C14, C61, D11, D61, J13, J22

Keywords: actionable sets, policy packages, partial identification, childcare, labor supply, option value

Yonsei University. Email: eunixon@yonsei.ac.kr. This draft is circulated for discussion; comments and suggestions are welcome.

Appendix A. Duality and Collapse Certificates

A.1. Minimum enablement

Write the primal program as

$$\min_{a\in\mathbb{R}^d,\, e\in\mathbb{R}^m_+} p'e \quad \text{s.t.} \quad -Ba + He \geq -b.$$

The action vector is free and the enabling vector is componentwise nonnegative.

Proof of Proposition 1. Let $\lambda \geq 0$ satisfy $B'\lambda = 0$ and $H'\lambda \leq p$. For every primal-feasible (a, e),

$$-\lambda' b \leq \lambda' He \leq p'e.$$

The first inequality follows by multiplying $Ba - He \leq b$ by λ and using $B'\lambda = 0$; the second follows from $H'\lambda \leq p$ and $e \geq 0$. This gives weak duality. The primal objective is bounded below by zero. Feasibility and finite-dimensional linear-program duality give primal and dual optimizers and equality of their values.

Complementary slackness gives $\lambda^{*\prime}(Ba^* - He^* - b) = 0$ and $e_r^*(p_r - H_r'\lambda^*) = 0$ for every input r. Since $B'\lambda^* = 0$, the first equality becomes $-\lambda^{*\prime}b = \lambda^{*\prime}He^*$. The second gives $\lambda^{*\prime}He^* = p'e^*$. For every used input, $H_r'\lambda^* = p_r$. □

If inputs are grouped by provider, $H = (H_1, \ldots, H_K)$ and $e = (e_1, \ldots, e_K)$. The accounting shares satisfy

$$\sum_{k=1}^{K} \lambda^{*\prime} H_k e_k^* = \lambda^{*\prime} He^* = C^*.$$

The equality permits several optimal share vectors when the dual solution is set-valued. The empirical tables report lower and upper shares over the optimal dual face in that case.

A.2. Farkas and Helly certificates

For the current system $Ba \leq b$, Farkas' lemma gives exactly one of the following two systems a solution:

$$Ba \leq b,$$

or

$$\lambda \geq 0, \qquad B'\lambda = 0, \qquad \lambda' b < 0.$$

The second system is a collapse certificate. Its normalization can be fixed by $-\lambda' b = 1$, $\|\lambda\|_1 = 1$, or the dual technology restriction $H'\lambda \leq p$, according to the reported object.

Let $K_1, \ldots, K_L$ be convex subsets of $\mathbb{R}^d$ generated by the action constraints. If $\cap_{\ell=1}^{L} K_\ell = \varnothing$, Helly's theorem gives $J \subseteq \{1, \ldots, L\}$ with $|J| \leq d + 1$ and $\cap_{\ell \in J} K_\ell = \varnothing$. The empirical certificate reports the selected rows, their economic labels, and the action-space dimension. Enumeration across discrete job cells precedes the convex certificate calculation.

A.3. Value-function derivatives

Let $\Lambda^*(b, p, H)$ and $E^*(b, p, H)$ denote the optimal dual and enabling sets. Directional derivatives follow the standard linear-program value formula. With unique solutions,

$$dC = -\lambda^{*\prime} db + e^{*\prime} dp - \lambda^{*\prime} (dH) e^* .$$

With set-valued solutions, the directional derivative is the corresponding extremum over the primal-dual optimal face. The implementation returns these bounds alongside the point estimate.

Appendix B. Repair-Set Proofs and Gate Representation

B.1. Minimal-path surfaces

PROOF OF THEOREM 1. Every minimal-path representation is nonnegative. If $S \subseteq S'$, then $B_{S,\mathcal{D}} \leq B_{S',\mathcal{D}}$ for every antichain, so $M(S) \leq M(S')$. The total represented mass bounds $M(S')$ by $\bar{M}$.

For the converse, list the distinct positive values of an isotone surface as $0 < v_1 < \cdots < v_L$ and set $v_0 = 0$. For each level define the upper set $\mathcal{U}_\ell = \{S : M(S) \geq v_\ell\}$. Its inclusion-minimal members form an antichain $\mathcal{D}_\ell$, and $S \in \mathcal{U}_\ell$ exactly when $D \subseteq S$ for some $D \in \mathcal{D}_\ell$. Assign mass $v_\ell - v_{\ell-1}$ to $\mathcal{D}_\ell$. Then

$$\sum_{\ell=1}^{L}(v_\ell - v_{\ell-1})B_{S,\mathcal{D}_\ell} = M(S),$$

and total represented mass equals $M(\mathcal{T}) \leq \bar{M}$. Proposition 2 and Möbius inversion give the singleton-antichain statement, including uniqueness. □

B.2. Exact inversion

PROOF OF PROPOSITION 2. For any package S, partition agent–option pairs by their exact repair set. The activation rule gives

$$M(S) = \mathbb{E}\left[\sum_j w_{ij} \sum_{T \subseteq \mathcal{T}} \mathbb{1}\{D_{ij} = T\}\mathbb{1}\{T \subseteq S\}\right] = \sum_{T \subseteq S} \theta(T).$$

Apply Möbius inversion on the subset lattice to obtain

$$\sum_{S \subseteq T}(-1)^{|T|-|S|}M(S) = \sum_{R \subseteq T}\theta(R) \sum_{S : R \subseteq S \subseteq T}(-1)^{|T|-|S|} = \theta(T),$$

because the inner sum equals one for $R = T$ and zero for every proper subset. Nonnegativity follows from $w_{ij} \geq 0$. Summing exact masses over sets of cardinality r gives Θ_r. The displayed inversion is the saturated contrast over arms contained in T. □

B.3. Package-surface characterization

PROOF OF THEOREM 2. Proposition 2 gives the equivalence between statements (i) and (ii), including uniqueness. Suppose statement (i) holds and let $R \cap S = \varnothing$. Substitution into the mixed increment gives

$$\Delta_R M(S) = \sum_{Q \subseteq R}(-1)^{|R|-|Q|} \sum_{T \subseteq S \cup Q} \theta(T).$$

For a fixed $T \subseteq S \cup R$, the coefficient on $\theta(T)$ equals one when $R \subseteq T$ and zero otherwise. Every retained set has the form $T = R \cup V$ for $V \subseteq S$, so $\Delta_R M(S) = \sum_{V \subseteq S} \theta(R \cup V) \geq 0$. This proves statement (iii). Conversely, statement (iii) evaluated at $S = \varnothing$ gives $\Delta_T M(\varnothing) \geq 0$ for every T, and this quantity is the Möbius coefficient in statement (ii). For $T = \varnothing$, the coefficient is $M(\varnothing)$. Hence statement (iii) implies statement (ii). □

B.4. Limited support

Proof of Theorem 3. Every repair distribution compatible with the model is nonnegative, has total mass at most $\bar{M}$, and satisfies $A_{\mathcal{O}}\theta = m_{\mathcal{O}}$. Conversely, any vector satisfying these restrictions defines a population distribution over exact repair sets with the observed active-set masses: assign mass $\theta(T)$ to repair set T and assign the residual $\bar{M} - \mathbf{1}'\theta$ to options outside the declared input universe. Hence the displayed polytope contains exactly the observationally equivalent repair distributions. Linear minimization and maximization over this set attain the sharp bounds because the set is closed and bounded.

If $q = r'A_{\mathcal{O}}$, then $q'\theta = r'm_{\mathcal{O}}$ for every feasible θ, so the functional is point identified. Full column rank provides this representation for every q and identifies the vector. Interval measurements replace point restrictions with their declared admissible set; the same construction establishes sharpness of the expanded polytope. □

The path-distribution sharpness argument replaces repair-set labels by antichain labels and $A_{\mathcal{O}}$ by $B_{\mathcal{O}}$. Any feasible ψ assigns its mass to options whose minimal path collection equals the corresponding antichain. The residual candidate mass is assigned outside the declared input universe. This construction generates the supported package means and proves sharpness of $\mathcal{H}^P$.

For the arm-width criterion, any ordered pair in one fiber satisfies $a_S'(\theta^1 - \theta^2) = 0$, and every pair satisfying this equality belongs to the fiber indexed by θ^1. Maximizing the target difference over such pairs gives the displayed LP. The value equals zero exactly when every fiber assigns one value to $q'\theta$. If q belongs to the row space of the augmented matrix $[A_{\mathcal{O}}; a_S']$, supported means together with the added-arm mean determine $q'\theta$, so every fiber has zero width.

For the nested two-input design $\mathcal{O} = \{\varnothing, \{d\}, \{d, w\}\}$, the observed equations are

$$M(\varnothing) = \theta_{\varnothing}, \qquad M(\{d\}) = \theta_{\varnothing} + \theta_d, \qquad M(\{d, w\}) = \theta_{\varnothing} + \theta_d + \theta_w + \theta_{dw}.$$

Thus $\theta_d = M(\{d\}) - M(\varnothing)$ and $\theta_w + \theta_{dw} = M(\{d, w\}) - M(\{d\})$. Observing the arm $\{w\}$ adds $M(\{w\}) = \theta_{\varnothing} + \theta_w$ and separates the remaining pair.

B.5. Threshold gates

Option j uses inputs in R_j. Input k satisfies its requirement when $x_{jk}(e,\omega) = \mathbb{1}\{e_k \geq \tau_{jk}(\omega)\}$. A monotone gate ϕ_j maps the status vector into activation, and $\Delta_j(\omega) \geq 0$ supplies an option weight. Define $V_j(e) = \mathbb{E}[\Delta_j \phi_j(x_j(e,\omega))]$. For coordinates $r \neq s$, let

$$\delta_{rs}\phi_j(x_{-rs}) = \phi_j(1,1,x_{-rs}) - \phi_j(1,0,x_{-rs}) - \phi_j(0,1,x_{-rs}) + \phi_j(0,0,x_{-rs}).$$

LEMMA A1 (Gate cross-difference). *For increments $h_r, h_s > 0$,*

$$\nabla_{rs}^{h} V_j(e) = \mathbb{E}\Big[\Delta_j \mathbb{1}\{e_r < \tau_{jr} \leq e_r + h_r,\ e_s < \tau_{js} \leq e_s + h_s\}\delta_{rs}\phi_j(x_{j,-rs}(e,\omega))\Big].$$

The conjunction gate has a weakly positive cross-difference, and the disjunction gate has a weakly negative cross-difference.

PROOF. For a realized state, coordinate r changes status across its increment exactly on $\{e_r < \tau_{jr} \leq e_r + h_r\}$, with the analogous event for s. The four-term input contrast vanishes unless both events occur. On their intersection it equals $\Delta_j \delta_{rs}\phi_j(x_{j,-rs})$. Integration gives the formula. The Boolean cross-difference equals one for a conjunction when every remaining required coordinate is satisfied. It equals minus one for a disjunction when every remaining coordinate is unsatisfied. □

The exact-repair model corresponds to a conjunction over a unique missing set at the declared policy doses. A gate with several activation paths generates an antichain in Theorem 1. Package monotonicity evaluates the path model, and Möbius nonnegativity evaluates its singleton-antichain restriction.

Appendix C. Welfare and Identification

C.1. Projection inference

Proof of Corollary 1. Let $m_0 = A_{\mathcal{O}}\theta_0$ denote the true supported active-set masses. On the event $m_0 \in \mathcal{C}_M(1-\alpha)$, the definition of $\mathcal{C}_\theta(1-\alpha)$ contains θ_0. This event has probability at least $1-\alpha$. Every declared linear functional evaluated at θ_0 lies inside its projection over the same set on that event, which gives simultaneous coverage. □

C.2. Choice-equivalence proof

Proof of Proposition 3. For the first system, let utility type a assign a strictly higher utility to action a than to every other action, and draw type a with probability $p(a \mid x)$. Universal availability then generates $p(\cdot \mid x)$. For the second system, fix $u(j) > u(0)$ for each $j \in \mathcal{J}$. Draw active set $\{j\}$ with probability $p(j \mid x)$ and the empty set with probability $p(0 \mid x)$. Choice from the active set together with the outside action again generates $p(\cdot \mid x)$. Universal availability and stochastic singleton availability have different active-set distributions. Applying the construction at each value of x, including each assignment arm, gives the conditional statement. □

C.3. Repair-targeting proof

Proof of Proposition 4. Every cell in $\mathcal{G}^R_{iT}$ has repair set T. Under package S, such a cell belongs to the active set exactly when $T \subseteq S$. For $T \nsubseteq S$, choice from the active set assigns probability zero to $\mathcal{G}^R_{iT}$, which gives the first result. The event $\{C_i(T) \in \mathcal{G}^R_{iT}\}$ is contained in $\{\mathcal{G}^R_{iT} \neq \varnothing\}$, which gives the bounds. Random assignment equates each observed arm mean with the corresponding population response probability. □

C.4. Envelope proof

Proof of Proposition 5. For a realized state, write the Lagrangian as

$$\mathcal{L}_i(a, \lambda; e_i, \omega) = u_i(a, \omega) - \lambda' g_i(a, q_i, \omega; e_i).$$

The constraint qualification, uniqueness, and envelope theorem give

$$\frac{\partial v_i(e_i, \omega)}{\partial e_{ki}} = -\lambda_i(e_i, \omega)' \frac{\partial g_i(a_i^*, q_i, \omega; e_i)}{\partial e_{ki}}.$$

The domination condition permits differentiation through the expectation. If the input weakly lowers each affected constraint function, the inner product has the stated sign. □

At threshold states, classical differentiability can fail. Directional welfare changes are then computed from the finite menu gain. The continuous multiplier formula and discrete activation formula are reported on their respective domains.

C.5. Implementation proof

PROOF OF PROPOSITION 6. The provider's marginal return after the input payment is

$$\frac{\partial r_{ki}}{\partial e_{ki}} + \alpha_{ki}\frac{\partial W_i}{\partial e_{ki}} + s^*_{ki}.$$

Substituting the proposed rate gives $\partial r_{ki}/\partial e_{ki} + \partial W_i/\partial e_{ki} + \sum_{h\neq i}\partial W_h/\partial e_{ki}$, which is the planner's marginal return. Local concavity makes the shared first-order condition sufficient for local implementation. Substitution of the envelope formulas gives the multiplier expression. □

C.6. Transfer spanning

PROOF OF PROPOSITION 7. A transfer portfolio produces state profile $Z\theta$. Exact replication requires a solution to $Z\theta = \Delta_b$, which exists exactly when Δ_b belongs to the column space of Z. In a finite state space, column-space membership is equivalent to $\operatorname{rank}(Z) = \operatorname{rank}([Z, \Delta_b])$. □

For approximate replication under state weights Q, the residual is

$$R_b^2 = \min_\theta (\Delta_b - Z\theta)' Q(\Delta_b - Z\theta).$$

The instrument class, state partition, and weight matrix are fixed before this residual is interpreted.

C.7. Outside-choice equivalence

Let the observed action space be $\{0, 1\}$. In environment A, option 1 is active and $u(0) > u(1)$. In environment B, the active set is empty. Both environments produce action 0 with probability one. Repeated choice under the same menu and state retains this equivalence. Availability shifters, utility shifters, or auxiliary constraint measurements supply additional variation.

Suppose z_1^U enters only $u(1)$, and its support contains a value at which $u(1) > u(0)$ whenever option 1 is active. Choice of option 1 at some support point reveals activity. Persistent choice of 0 over the support gives an inactivity conclusion under the stated utility support condition. With bounded support, the conclusion becomes set-valued.

Appendix D. Data Construction

D.1. Households and job cells

The household sample contains adults with children in the policy-relevant age range. Household links assign children, partners, education, income, vehicle access, and work outcomes to each adult. Job cells are estimated from workers outside the outcome sample by state, PUMA, occupation, hours band, departure-time band, commute band, and remote-work status. Sparse cells are pooled using a rule fixed before outcome estimation. The national construction combines PUMA occupation weights with state occupation-specific schedule and commute distributions.

Let t_j^- and t_j^+ denote a job cell's work interval and let r_{ij}^-, r_{ij}^+ denote outward and return travel time. Let $c_i^-(d)$ and $c_i^+(d)$ denote the care window under regime d. The care-commute constraint is

$$c_i^-(d) \le t_j^- - r_{ij}^-, \qquad t_j^+ + r_{ij}^+ \le c_i^+(d).$$

The lower compatibility set uses the upper travel-time bound and the inner care-window bound. The upper compatibility set uses the lower travel-time bound and the outer care-window bound. Vacancies, employer acceptance, available care slots, and prices enter a later actionability classification.

D.2. Prices and resource bounds

Required weekly paid care is the uncovered duration between the job-travel interval and the public or household-provided care window. County-age-year prices map this duration into expenditure. The resource bound uses observed income and program copay rules. Cash-transfer comparisons alter this bound while holding the coverage window fixed; duration comparisons alter the window while holding the declared price schedule fixed.

D.3. Policy-event record

Every event record contains the legal or administrative source, announcement date, effective date, eligible ages, daily hours, covered days, geographic unit, implementation share, tuition or copay rule, and concurrent funding changes. Event inclusion requires a coverage or enrollment first stage in the same geographic unit. The source ledger stores snapshots and hashes before outcome estimation.

The ledger transcribes mandate and full-day-friendly policy dates for the 50 states and the District of Columbia from Table C1 of the published supplement to *A Matter of Time?*. The acquired OpenICPSR V1 archive supplies the corresponding policy code and CPS construction files. The extension cross-checks the two artifacts and selects Connecticut, Hawaii, Indiana, Kansas, Massachusetts, Montana, North Dakota, Ohio, Oregon, Rhode Island, Washington, and Wyoming, whose first policy dates fall from 2007 through 2017. The published supplement describes its state-code, Education Commission of the States, National Conference of State Legislatures, and media review. The archive omits the underlying state-specific legal snapshots, so the extension records the calendar cross-check and leaves the two-record legal-source criterion open.

The comparison pool contains Alaska, Arizona, California, Florida, Iowa, Maine, Michigan, Missouri, Nebraska, Pennsylvania, South Dakota, Texas, Vermont, and Wisconsin. The outcome window is school years 2001–2022. The CPS October Supplement supplies the public-school full-day kindergarten share for children ages three through seven; a centered three-year mean follows the source construction. March CPS ASEC records in calendar year (t+1) receive the October measure from school year (t). Mothers with an own child age five or six form the exposed group, while mothers whose own children in the five-to-nine range are ages seven through nine form the within-market comparison.

The frozen map uses 482,285 workers ages 25–54 from 2005 ACS one-year PUMS files for the 26 analysis states. State, two-digit occupation, and weekly-hours bands define 10,996 cells. Departure time, commute, remote status, and weekly hours determine whether each worker's imputed interval fits 8 a.m.–noon or 8 a.m.–3 p.m. A cell's repair score is the 2005 weighted share that fits the longer window and lies outside the shorter window. State cells with fewer than 30 records use the pooled occupation–hours score. The predicted gain set begins at the worker-weighted seventy-fifth percentile of positive scores, 0.268. This map classifies 99.7 percent of weighted employed mothers in the CPS outcome sample. All 80 ACS replicate-weight maps use the same cell keys, pooling rule, and cutoff.

D.4. National benchmark implementation

The 2024 national PUMS person archive contains 3,422,888 records across two CSV members. The target sample contains 112,533 mothers and 12,599,726 weighted people. Reference workers exclude this target group and supply state cells with at least five records. Two-digit occupation, weekly-hours, departure-time, commute, and remote-work categories define a cell. The weighted tenth percentile of schooling supplies its credential threshold. Remote work sets commute to zero and uses an 8 a.m. start. Daily work duration equals usual weekly hours divided by five; the SIPP validation replaces this divisor with reported days worked and compares the implied duration with reported start and end bins.

Within each state and occupation, the reference-cell weights define a schedule and commute distribution. PUMA reference workers define the local occupation weights. Their product gives each PUMA's model-assisted job-cell distribution. A PUMA enters when its retained occupations contain at least 150 reference workers across at least 10 occupation groups. State cells supply the distribution for other PUMAs. The PUMA rule covers 86.7 percent of weighted target mothers, with regional coverage from 82.2 percent in the Midwest to 92.4 percent in the West.

For each education level, the denominator is total local reference-worker cell weight. The numerator adds cells whose credential threshold is satisfied and whose departure-to-return interval lies inside the care window. Person-weighted means aggregate the resulting education-specific compatibility masses across target mothers. Joint intervals use all 80 ACS replicate weights. Each replicate reweights retained state job cells, PUMA occupation composition, and target households; schedule bands, credential thresholds, repair labels, and the PUMA support rule are frozen before reweighting. The ACS factor 4/80 forms the design covariance matrix, and a Gaussian max-t calculation gives simultaneous bands across the package surface and exact repair vector. Pooling sensitivity records variation from alternative local-support thresholds.

D.5. Held-out SIPP validation

The 2024 SIPP public-use file covers January through December 2023. The validation sample selects women ages 25–54 identified as reference parents in households containing a child under age six. SIPP job records report days worked per week, typical start and end time bins, schedule type, and work-from-home days. Childcare records report center and preschool use, payments, waiting lists, and whether arrangements prevented the reference parent from working or working more.

For a time bin $[\underline{t}, \overline{t}]$, the lower compatibility indicator requires $\underline{t}^{\text{start}}$ to follow the care opening and $\overline{t}^{\text{end}}$ to precede the care closing. The upper indicator requires some pair of times in the two reported bins to fit the window. The same file evaluates the ACS daily-hours proxy by comparing weekly hours divided by reported days with the midpoint duration between reported start and end bins.

D.6. Sample separation

Job-cell weights, schedule distributions, commute mappings, and credential rules come from reference workers outside the target group. Target mothers merge after cell construction by PUMA or state and schooling level. The SIPP benchmark uses a separate survey and calendar year.

D.7. Uganda replication analysis

The Uganda archive contains 1,496 randomized households: 412 control, 363 childcare, 364 cash, and 357 combined-arm households. The experimental protocol offered one target child one year of full-day care at a nearby registered center, assisted enrollment, and

TABLE A1. Held-out SIPP work-interval and childcare-friction benchmark

Work-interval test	Lower bound	Upper bound
8 a.m.–noon	0.022	0.029
8 a.m.–3 p.m.	0.093	0.134
7:30 a.m.–5:30 p.m.	0.542	0.610
Full-day class	**Care limits work**	**Care waitlist**
Compatible	0.107	0.114
Boundary interval	0.044	0.077
Outside window	0.064	0.058

Notes: The sample uses 2024 SIPP records covering calendar year 2023. Work-interval lower and upper bounds account for the reported start and end time bins. Care limits work equals one when the respondent reports that childcare arrangements prevented work or additional work. Care waitlist equals one when the respondent reports a childcare waiting list. Person weights produce all shares. Employment and care arrangements are selected in these cross-sectional groups, so the final two columns serve as measurement diagnostics.

paid the center each trimester. Target-child identifiers link the baseline, short-term, and long-term household rosters. The reconstruction follows the authors' code by replacing missing short-term enrollment with retrospective 2019 enrollment and replacing missing short-term program status with the long-term program field. Program code 1 enters half-day care and code 2 enters full-day care. ANCOVA specifications include treatment-arm indicators, randomization strata, the baseline outcome, a baseline-missing indicator, and HC1 standard errors.

Entry into any care is evaluated among children unenrolled at baseline, and entry into full-day care is evaluated among children outside full-day care at baseline. These transitions record realized program use. The assigned childcare factor averages the care effect across randomized cash states. The younger-sibling interaction compares this factor effect across the two baseline repair classes. The archive contains treatment assignment, enrollment, program duration, household rosters, labor records, and business records. The job-level actionable set, work-schedule compatibility, transport compatibility, and care arrangements for the younger sibling remain outside the measured fields. The source assessment therefore assigns the archive to randomized input delivery, use, incomplete-package response, and transition diagnostics.

The labor reconstruction follows the original preparation code. Wage hours sum reported days times daily hours across the respondent's jobs and cap monthly hours at 720. Self-employment hours combine owner and co-owner records and apply the same cap. Wage earnings and business profits are summed across the respondent's records, expressed in thousands of Ugandan shillings, and winsorized at the long-term 99th percentile. Interviewed households with no reported wage job or business receive zero for the corresponding outcome. The resulting 1,414-observation regressions reproduce every childcare coefficient, younger-sibling effect, and childcare–younger-sibling interaction in published Table A.9 within 0.005 of its reported two-decimal value.

D.8. UNPS target and transport construction

The population target uses the 2019–2020 Uganda National Panel Survey household identification, roster, and education files, named `gsec1`, `gsec2`, and `gsec4` in the source archive. The source archive requires registration and remains outside the replication package. A manifest records its DOI, catalog identifier, byte count, member count, CRC result, and SHA-256 hash. The code reads the three files directly from the user-acquired archive and writes aggregate estimates, balance records, and source metadata. Generated outputs contain aggregate results and source summaries.

The initial target contains 171 children ages four through six in Masaka, Mityana, Mukono, Iganga, Mbale, Jinja, Kabarole, Kasese, and Kyenjojo. Subtracting one from survey age follows the population comparison in the trial replication code. Household roster counts provide size, and the presence of a different household member with a lower age defines the common repair class. The support rule retains household sizes two through thirteen, the observed range in the trial common-variable sample. Ten children in three larger households fall outside this range; their survey weights form 3.66 percent of initial target weight. The resulting aggregate target contains 161 children in 137 households. The no-younger-member class also excludes aligned age five because that class–age cell is absent from the trial.

The target and trial basis is declared before model fitting. Age four, age five, districts two through nine, female sex, current school attendance, the younger-member class, scaled household size, and its square supply fifteen columns. The selection model, four arm-specific outcome models, and four arm-specific outcome-observation models use logistic links with fixed ridge constants. Fold assignment occurs at village level in the trial and household level in the survey. Cross-fitted selection odds receive an exponential calibration tilt that equates every nonconstant target moment. No basis search or rank choice uses an outcome realization.

TABLE A2. Trial-to-target covariate balance

Covariate	Trial	Target	Weighted trial	Weighted SMD
Age four	0.410	0.347	0.347	0.000
Age five	0.097	0.249	0.249	0.000
Mityana	0.111	0.086	0.086	0.000
Mukono	0.098	0.161	0.161	0.000
Iganga	0.140	0.050	0.050	0.000
Mbale	0.126	0.168	0.168	0.000
Jinja	0.134	0.154	0.154	0.000
Kabarole	0.108	0.083	0.083	0.000
Kasese	0.069	0.139	0.139	0.000
Kyenjojo	0.102	0.043	0.043	0.000
Female child	0.496	0.544	0.544	-0.000
Currently attending school	0.361	0.822	0.822	-0.000
Younger household member	0.347	0.549	0.549	-0.000
Household size, scaled	-0.243	0.174	0.174	0.000
Household size squared, scaled	0.528	0.477	0.477	-0.000

Notes: Target means use UNPS survey weights over the supported aggregate target. Weighted-trial means use the calibrated density ratio. SMD divides the weighted mean difference by the square root of the average trial and target variance. District indicators omit Masaka and age indicators omit aligned age three.

Appendix E. Additional Results

E.1. Replication outputs

The replication package contains the official 2024 ACS national PUMS person archive and source hash, state schedule cells, PUMA occupation weights, household compatibility maps, the 26-window frontier, timing-block repair masses, nested-package bounds, sampling intervals, a chosen-cell check, and the held-out 2024 SIPP benchmark. The analysis record stores sample counts, assumptions, output hashes, regional estimates, the coverage grid, and pooling sensitivity.

E.2. PUMA pooling sensitivity

TABLE A3. PUMA pooling sensitivity

Specification	PUMA share	Half-day	Full-day	Extended	Chosen–local
State cells	0.000	0.008	0.023	0.235	0.071
PUMA 250	0.506	0.008	0.023	0.238	0.071
PUMA 150	0.867	0.008	0.024	0.237	0.071
PUMA 100	0.967	0.008	0.024	0.237	0.071

Notes: Each row changes the minimum reference-worker count for PUMA occupation weights while retaining the 10-occupation requirement. State cells supply local composition outside supported PUMAs. Mass columns report national person-weighted compatibility shares. Chosen–local subtracts household-specific full-day compatibility mass from observed chosen-job compatibility among employed target mothers.

Across the four rules, half-day mass ranges from 0.81 to 0.83 percent, full-day mass ranges from 2.31 to 2.39 percent, extended-day mass ranges from 23.49 to 23.80 percent, and the chosen–local gap ranges from 7.06 to 7.13 percentage points.

E.3. Coverage-grid identities

For each geography and opening time, compatibility mass rises weakly over the 13 closing times from noon through 6 p.m. At every closing time, the 7:30 a.m. opening admits weakly more mass than the 8 a.m. opening. The national 8 a.m. frontier reproduces the half-day and full-day estimates at noon and 3 p.m., and the 7:30 a.m. frontier reproduces extended-day mass at 5:30 p.m.

The seven reported repair types sum to full-package mass in every Census region and the national sample. A continuous departure-to-return interval that requires both the early and late blocks also requires the midday block, so the $\{e, l\}$ cell equals zero. The replication test checks this restriction to machine precision together with the package adding-up identities.

E.4. Joint reference–target inference

The ACS one-year file supplies 80 replicate weights for every reference worker and target mother. Each replicate rebuilds retained state-cell weights, PUMA occupation composition, and target aggregation while keeping cell definitions and repair labels fixed. The base replicate rebuilds all eight package means and exact repair masses with maximum error 3.4×10^{-16}. The design covariance uses the ACS factor 4/80. A Gaussian max-t calculation across the eight repair coefficients gives critical value 2.632. The full-package interval is [23.55, 23.93] percentage points; the midday-and-late repair interval is [17.03, 17.35], and the all-three-block interval is [3.87, 4.03].

E.5. Uganda microdata reconstruction

TABLE A4. Childcare use and transitions in the Uganda replication data

Outcome	Control	Childcare	Cash	Combined	Joint	p	N
Any childcare use, short term	0.823	0.149 (0.021)	0.074 (0.025)	0.138 (0.022)	-0.085 (0.029)	0.003	1,428
Full-day childcare use, short term	0.373	0.457 (0.032)	0.051 (0.035)	0.477 (0.031)	-0.032 (0.045)	0.482	1,430
Entry into any childcare, short term	0.719	0.239 (0.032)	0.122 (0.038)	0.228 (0.033)	-0.133 (0.043)	0.002	910
Entry into full-day childcare, short term	0.363	0.460 (0.032)	0.049 (0.036)	0.479 (0.031)	-0.030 (0.046)	0.514	1,404
Full-day childcare use, long term	0.492	0.349 (0.033)	0.089 (0.036)	0.378 (0.032)	-0.059 (0.046)	0.192	1,358

Notes: Each arm column reports an ANCOVA coefficient relative to control, with HC1 standard errors in parentheses. Regressions include the authors' district, younger-sibling, baseline-schooling, caregiver-occupation, and caregiver-relation strata controls. Use outcomes add the baseline outcome and a missing-baseline indicator. Entry samples condition on absence of the stated care form at baseline. Joint is the combined-arm coefficient minus the childcare and cash coefficients; the adjacent column reports its two-sided p-value. Enrollment and program duration measure realized use.

The childcare arm raises entry into any care by 0.239 and entry into full-day care by 0.460. The cash arm raises the same transitions by 0.122 and 0.049. The combined-arm coefficients are 0.228 and 0.479. The joint contrast for any-care entry is −0.133, with p-value 0.002; the corresponding full-day contrast is −0.030, with p-value 0.514. Long-term full-day use rises by 0.349, 0.089, and 0.378 across the childcare, cash, and combined arms. These quantities characterize delivered use and behavioral transitions under randomized assignment.

The younger-sibling analysis uses the same baseline stratum as the original heterogeneity specification. The target-child use regressions contain 1,430 observations at short term, 1,404 for short-term full-day entry, and 1,358 at long term. The economic-response regressions contain all 1,414 long-term household interviews. The reconstruction reproduces published Online Appendix Table A.9: the largest absolute gap between a reconstructed childcare-arm coefficient, its younger-sibling effect, or their interaction and the corresponding two-decimal entry is 0.0049. The reported main-text factor contrast averages the childcare marginal effect over the two randomized cash states.

E.6. Randomized supply-path diagnostic

The childcare and cash assignments form a two-input package surface for target-child use. Short-term full-day use, short-term full-day entry, and long-term full-day use rise weakly along every observed package inclusion at their point estimates. Their single-path joint coefficients are −3.17, −2.99, and −5.94 percentage points. The two-input antichain equations bound mass reachable through either singleton path by [3.17, 5.15], [2.99, 4.86], and [5.94, 8.86] points. Any-care use and entry have joint coefficients of −8.53 and −13.26 points, with two-sided p-values 0.003 and 0.002. Their smallest monotonicity gaps are −1.08 and −1.10 points, with lower-tail p-values 0.232 and 0.291. These calculations report the path restriction and the single-path restriction separately.

E.7. Uganda repair-class sensitivity

The negative profit and total-income contrasts persist across every upper-tail rule and district omission. Their magnitudes grow as strata and baseline controls enter the regression. The stratified raw randomization p-values are 0.035 and 0.053. Randomization max-t adjustment gives 0.193 and 0.311. The covariance-weighted randomization Wald gives p-value 0.017, and the sum-of-squared-t omnibus gives 0.187. Cash heterogeneity is 3.26 for profits and 3.90 for total income; subtracting it from childcare heterogeneity gives −18.76 and −21.11. The two omnibus statistics weight the outcome vector differently. Component adjustments report the precision of each outcome contrast within the nine-outcome family.

The joint contrasts are negative for the four labor-supply outcomes, with additivity p-values of 0.018 for self-employment, 0.038 for self-employment hours, 0.013 for employment, and 0.073 for total hours. An exact repair distribution assigns nonnegative mass to every repair type. These outcome contrasts therefore combine activation with choice, substitution, and household allocation. The cash arm's enrollment effects also show that delivered care can move inside an assignment cell whose program label names another input.

TABLE A5. Sensitivity of the Uganda repair-class contrast

Specification	Self-employment profits	Total labor income
Panel A. Upper-tail treatment with strata and baseline controls		
None	−17.92 (8.99); p = 0.046	−19.21 (10.37); p = 0.064
99.5th percentile	−16.72 (8.11); p = 0.039	−17.92 (10.12); p = 0.077
99th percentile	−15.50 (7.29); p = 0.034	−17.21 (9.16); p = 0.060
97.5th percentile	−12.06 (6.02); p = 0.045	−15.35 (8.03); p = 0.056
95th percentile	−9.28 (5.04); p = 0.065	−11.01 (6.76); p = 0.104
Panel B. Covariate specification at the published 99th-percentile rule		
Assignment cells	−11.88 (7.67); p = 0.122	−13.07 (9.67); p = 0.177
Assignment cells and strata	−13.39 (7.36); p = 0.069	−13.58 (9.36); p = 0.147
Assignment cells, strata, and baseline outcome	−15.50 (7.29); p = 0.034	−17.21 (9.16); p = 0.060
Panel C. Influence, multiplicity, and treatment-component contrasts		
Leave-one-district estimate range	-19.11 to -11.23	-24.47 to -13.62
Leave-one-district p-value range	0.016 to 0.108	0.011 to 0.139
Holm-adjusted p-value across nine outcomes	0.303	0.484
Stratified randomization-inference p-value	0.035	0.053
Randomization max-t adjusted p-value	0.193	0.311
Childcare minus cash heterogeneity	−18.76 (10.81); p = 0.083	−21.11 (13.36); p = 0.114
Joint test: nine outcome interactions	20.14 on 9 df; p = 0.017	
Randomization omnibus: nine outcomes	sum of squared t = 13.78; p = 0.187	
Randomization Wald: nine outcomes	20.14; p = 0.017	

Notes: Each regression estimates the younger-sibling difference in the average marginal childcare effect across randomized cash states. Cells in Panels A and B report the interaction, HC1 standard error in parentheses, and two-sided p-value. Monetary outcomes are thousands of Ugandan shillings in the prior month. Panel A changes the upper-tail cap and retains randomization strata and the baseline outcome. Panel B changes the control set under the source code's 99th-percentile rule. Panel C omits each of nine districts in turn, applies Holm and randomization max-t adjustments to the published nine-outcome family, uses 4,999 treatment permutations within 156 randomization strata, compares childcare and cash heterogeneity, and reports sum-of-squared-t and covariance-weighted Wald omnibus tests.

E.8. Public-school duration design assessment

The extension specification freezes its cohort, comparison states, job-cell map, outcome family, event window, equivalence bands, and inference procedure before outcome estimation. The analysis contains 134,917 mothers across 26 states and school years 2001–2021. State-specific metropolitan-area-by-year effects absorb the policy main effect and local labor-market conditions, while state-by-exposure-group effects absorb persistent differences between the child-age groups. Standard errors cluster by state. A common state-level Rademacher score bootstrap and the 80 ACS replicate maps form the confirmatory max-(t) distribution.

The policy coefficient raises full-day enrollment by 0.226. Its simultaneous interval excludes zero. Employment changes by 0.003, full-time employment by 0.006, gain-set employment by 0.001, and interstate migration by 0.002. Every labor and migration simultaneous interval contains zero. The gain-set combined standard error is 0.010, of which

TABLE A6. Factorial outcome and delivery sign diagnostics

Outcome	Childcare	Cash	Combined	Joint contrast	*p*-value
Self-employed	0.13	0.20	0.12	-0.21	0.018
Self-employment hours	35.68	52.90	28.72	-59.86	0.038
Employed	0.14	0.14	0.06	-0.22	0.013
Total working hours	19.24	34.97	0.78	-53.43	0.073
Any care, lower bound	0.14	0.07	0.13	-0.08	0.003
Full-day care, lower bound	0.48	0.06	0.50	-0.04	0.258

Notes: Entries in the first four rows are treatment coefficients for single mothers reported by Bjorvatn et al. (2025, Online Appendix Table A.8). The final two rows use lower-bound childcare-enrollment estimates from Online Appendix Table B.1, Panel A. Joint equals the combined-arm coefficient minus the childcare and cash coefficients. The final column reports the published p-value for equality of the combined effect and the sum of the two singleton effects. Employment and enrollment outcomes are indicators; hours are monthly.

0.006 comes from the 2005 ACS map. The employment first-stage ratio is 0.013.

The simultaneous lead endpoints reach 0.148 in absolute value for enrollment, 0.116 for employment, 0.084 for gain-set employment, and 0.041 for migration. Each value exceeds its registered equivalence band. The joint lead p-values are below 0.001 for enrollment and migration, 0.012 for employment, and 0.411 for gain-set employment. The event sample therefore fails the registered causal-transition gate. The paper uses the exercise to record delivered enrollment, mapping coverage, sampling uncertainty, and the reason this policy comparison is excluded from the causal evidence.

E.9. Transport estimator and attrition sensitivity

For the aggregate target, the augmented, outcome-regression, and selection-weighted Hájek estimates of the short-term childcare factor are 0.429, 0.436, and 0.412. Their long-term counterparts are 0.204, 0.203, and 0.198. The supported no-younger-member target gives 0.542, 0.505, and 0.526 at short term. The younger-member target gives 0.220, 0.389, and 0.228, so this smaller class carries sensitivity to the outcome model; its calibrated trial effective sample size is 89.6 and its target survey-weight effective sample size is 37.0 households' worth of independent weight.

Four fold assignments and three fixed regularization settings give aggregate transported childcare-factor ranges of [0.409, 0.434] at short term and [0.192, 0.204] at long term. Fixed-nuisance deletion of each target household in turn gives ranges of [0.409, 0.441] and [0.196, 0.227]; deletion of the ten trial villages carrying the largest calibrated weight gives [0.411, 0.439] and [0.191, 0.212]. Survey-weight effective sample sizes are 63.2 for the aggregate target, 20.4 for the supported no-younger-member target, and 37.0 for the younger-member target. The largest household weight shares are 0.066, 0.177, and

TABLE A7. Public-school duration design gate

Outcome	Estimate	Combined SE	95% simultaneous interval	Mean
Full-day enrollment first stage	0.226	0.053	[0.107, 0.344]	0.346
Employed last year	0.003	0.008	[-0.014, 0.020]	0.695
Full-time last year	0.006	0.012	[-0.022, 0.034]	0.496
Employed in frozen gain set	0.001	0.010	[-0.022, 0.023]	0.087
Moved across states	0.002	0.002	[-0.003, 0.006]	0.017
Max-t critical value				2.261
Employment IV ratio				0.013

Notes: The regressor is the interaction between a full-day-friendly policy and an indicator for an own child age five or six. The first-stage outcome interacts that child-age indicator with the state full-day enrollment share. Labor outcomes compare mothers with a child age five or six with mothers whose children in the five-to-nine range are ages seven through nine. The frozen gain set uses 2005 ACS state–occupation–hours cells. Combined standard errors add ACS replicate-map variation to state-cluster variation for the gain-set outcome. One max-(t) critical value covers the five displayed coefficients. The IV ratio divides the employment reduced form by the enrollment first stage.

0.075. The no-younger-member long-term point estimate ranges from 0.163 to 0.260 across the fitting specifications, so the repair-class estimates describe supported heterogeneity while the aggregate effect supplies the population headline.

Table A8 assigns every unobserved binary outcome to zero or one arm by arm. The bounds use trial weights for the experimental target and calibrated trial-to-target weights for the transported target. They are sample analogues under unrestricted outcome attrition and condition on the realized samples. The childcare-factor lower endpoint remains positive in every displayed sample. Long-term aggregate bounds are [0.209, 0.383] in the trial and [0.128, 0.245] in the supported target.

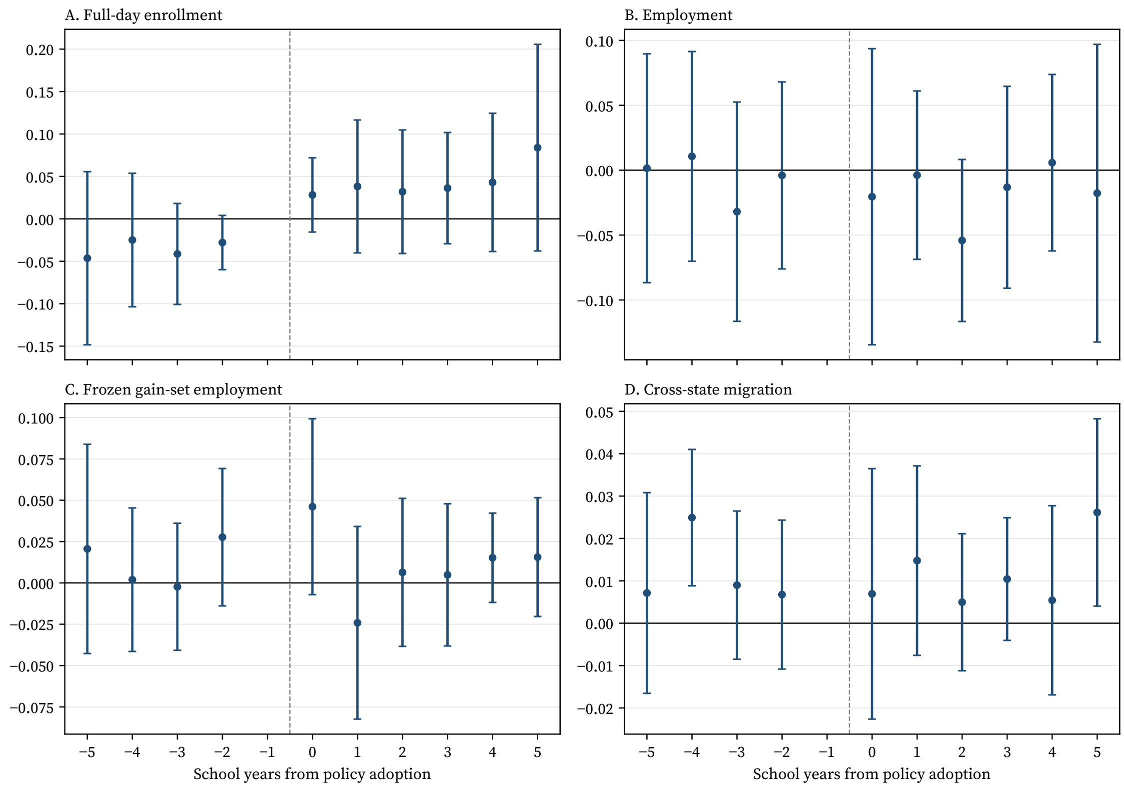

FIGURE A1. Event-time diagnostics for the public-school duration gate

Notes: Points are cohort-stacked event-time coefficients relative to school year -1. Whiskers give outcome-specific 95 percent simultaneous intervals from state-level Rademacher score draws. Each stack contains one treated state and the 14 comparison states. Panel C uses the 2005 frozen predicted gain set. The registered lead-equivalence bands are five percentage points for enrollment, two points for employment and gain-set employment, and one point for migration.

TABLE A8. Worst-case outcome-attrition bounds for the childcare factor

Outcome	Repair class	Trial	Nine-district target	Trial response	Target response
Full-day childcare use, short term	All target children	[0.383, 0.462]	[0.373, 0.426]	0.941–0.972	0.952–0.985
Full-day childcare use, short term	No younger household member (supported ages 3–4)	[0.404, 0.480]	[0.463, 0.548]	0.953–0.975	0.920–0.990
Full-day childcare use, short term	Younger household member	[0.346, 0.427]	[0.187, 0.253]	0.919–0.992	0.955–0.976
Full-day childcare use, long term	All target children	[0.209, 0.383]	[0.128, 0.245]	0.867–0.944	0.915–0.980
Full-day childcare use, long term	No younger household member (supported ages 3–4)	[0.239, 0.407]	[0.144, 0.283]	0.875–0.931	0.872–0.981
Full-day childcare use, long term	Younger household member	[0.151, 0.335]	[0.067, 0.218]	0.852–0.975	0.879–0.957

Notes: Brackets assign every missing binary outcome to zero or one separately in each randomized arm and then form the childcare-factor contrast. Response columns report the minimum and maximum arm-specific response rates under the population's weights. The intervals are identification-range sample analogues; the simultaneous sampling intervals in Table 6 provide the primary inferential statement.