

Labor Proxy Construction (Containers) — Terminal×Month Methodology (Clean Markdown)

Objective. Build a statistically careful **labor-hours proxy** $L_{\{i,m\}}$ at the **terminal × month** level for Israeli container terminals, using **published service KPIs** and **observed TEU throughput**. This report documents data, construction steps, assumptions, diagnostics, and robustness so it can be cited in the thesis methods.

1) Entities, Notation, Period, Units

- **Terminals (i):** Ashdod-Legacy, Ashdod-HCT (Southport), Haifa-Legacy, Haifa-Bayport (SIPG), Eilat (optional).
- **Ports (p):** Ashdod, Haifa, Eilat, with mapping $p(i)$.
- **Time:** Monthly m ; Year $y(m)$; Quarter $q(m)$.
- **Coverage:** Baseline **2018–2024** (exact TEU/hour available). Extension to **2014–2017** handled as an option (see §4).
- **Units:** TEU (containers); productivity **TEU per work-hour**; time metrics in **hours/ship**; responsiveness in **percent**.

Observed inputs (granularity → units)

- **Monthly port TEU:** $TEU_{\{p,m\}} \rightarrow \text{port} \times \text{month}$, TEU.
- **Quarterly terminal TEU:** $TEU_{\{i,q\}} \rightarrow \text{terminal} \times \text{quarter}$, TEU.
- **Annual TEU/hour (containers):** $Pi_{\{i,y\}} \rightarrow \text{terminal} \times \text{year}$, TEU/hour (published service KPI).

Quantities to construct

- $TEUhat_{\{i,m\}}$: **terminal×month** TEU (allocated from port months using observed quarterly shares).
 - $H_{\{i,y\}}$: **terminal×year** labor-hours.
 - $L_{\{i,m\}}$: **terminal×month** labor-hours (the labor proxy used in regressions).
-

2) Construction Steps (no smoothing across reforms)

Step 1 — Terminal shares (observed, by quarter)

For each port p and quarter q , compute the terminal's share of container TEU:

$$s_{\{i,p,q\}} = TEU_{\{i,q\}} / (\sum_{\{j \in I(p)\}} TEU_{\{j,q\}})$$

Structural zeros. Before a terminal opens (e.g., HCT pre-2022Q1; Bayport pre-2021Q3), set $TEU_{\{i, q\}}=0$ $\Rightarrow s_{\{i, p, q\}}=0$. These are **zeros**, not missing data.

Step 2 — Monthly TEU by terminal (allocation, not interpolation)

Allocate **port-monthly TEU** to terminals using the **quarter's** shares only:

$$TEUhat_{\{i, m\}} = s_{\{i, p(i), q(m)\}} \times TEU_{\{p(i), m\}}$$

This preserves **observed monthly seasonality at the port** and **observed intra-port shares at the quarter** without inventing within-quarter terminal swings.

Step 3 — Annual labor-hours via published TEU/hour

Aggregate terminal-monthly TEU to annual and divide by **published** TEU/hour:

$$\begin{aligned} TEU_{\{i, y\}} &= \sum_{\{m \in y\}} TEUhat_{\{i, m\}} \\ H_{\{i, y\}} &= TEU_{\{i, y\}} / Pi_{\{i, y\}} \end{aligned}$$

Use $Pi_{\{i, y\}}$ exactly **as published** (no smoothing). This respects step-changes from commissioning and privatization.

Step 4 — Monthly labor-hours (identity-preserving disaggregation)

Allocate $H_{\{i, y\}}$ to months by the terminal's **own** monthly TEU shares within the year:

$$\begin{aligned} w_{\{i, m|y\}} &= TEUhat_{\{i, m\}} / (\sum_{\{m' \in y\}} TEUhat_{\{i, m'\}}) \\ L_{\{i, m\}} &= H_{\{i, y(m)\}} \times w_{\{i, m|y(m)\}} \end{aligned}$$

This guarantees annual add-up: $\sum_{\{m \in y\}} L_{\{i, m\}} = H_{\{i, y\}}$. Pre-opening months remain 0 by construction.

3) Statistical & Econometric Validity

What we intentionally avoid

- **No within-quarter share interpolation.** Shares are piecewise-constant within a quarter and update only when a new quarter starts or a terminal opens.
- **No smoothing across structural breaks.** $Pi_{\{i, y\}}$ (TEU/hour) is used as published; we do **not** interpolate across HCT/Bayport commissioning or Haifa privatization.
- **No outcome leakage without checks.** L is derived from TEU and TEU/hour, so we will report robustness using an alternative labor proxy **L2** that does **not** scale by TEU (see §6).

Why the allocation is credible

- Relies only on **observed** monthly activity (port TEU) and **observed** intra-port competition (quarterly terminal shares). This is the **least-assumptions** route to terminal-monthly TEU.
- Avoids injecting high-frequency noise: within-quarter shares are constant.

Treatment timing is respected

- Commissioning appears as **structural zeros** → **positives** in $\text{TEUhat}_{\{i,m\}}$ and as real changes in $\text{Pi}_{\{i,y\}}$. No averaging dilutes events.

4) Edge Cases & Extensions

1) **2014–2017 backfill (optional)** - Preferred (no interpolation): $\text{Pi}_{\{i,y\}} \approx (\text{moves_per_crew_hour})_{\{i,y\}} \times (\text{TEU_per_box})_{\{p,y\}}$ using Yearbook container-unit counts. If units unavailable, either exclude 2014–2017 from the baseline or interpolate $\log \text{Pi}_{\{i,y\}}$ **within the pre-reform regime** (clearly flagged as sensitivity).

2) **Eilat** - Include if TEU/hour series exists; otherwise treat separately or omit from the baseline window.

3) **COVID & war months** - Flag months in 2020–2021 and late-2023/2024; do not smooth. Use flags as controls or window exclusions.

5) Diagnostics & QA (stored with the dataset)

- **Quarter reconciliation:** $\sum_{m \in q} \text{TEUhat}_{\{i,m\}} \stackrel{?}{=} \text{TEU}_{\{i,q\}}$
- **Annual reconciliation:** $\sum_{m \in y} \text{TEUhat}_{\{i,m\}} = \text{TEU}_{\{i,y\}}$ and $\sum_{m \in y} \text{L}_{\{i,m\}} = \text{H}_{\{i,y\}}$
- **Pre-opening zeros:** verified for HCT and Bayport.
- **Outlier flags:** months with extreme wait/berth/stay times are flagged (used only in robustness).
- **Versioning:** record KPI table version/hash and TEU panel version.

6) Robustness: Alternative Labor Proxy (L2)

Idea. Use **staffing intensity** rather than TEU/hour. - Annual person-teams: $\text{PT}_{\{i,y\}} = \text{Calls}_{\{i,y\}} \times \text{AvgTeamsPerShip}_{\{i,y\}}$ - Convert to **hours** by a single calibration constant per port-year so that: $\sum_m \text{L2}_{\{i,m\}} = \text{H}_{\{i,y\}}$ (keeps scale comparable; breaks the mechanical TEU link). - Monthly allocation uses $w_{\{i,m|y\}}$; if monthly calls ever exist, allocate by call shares.

Run core regressions with **L1** and **L2**; stability mitigates mechanical-correlation concerns.

7) Output Files & Schemas

- `labor_hours_monthly_terminal.tsv`

```
port terminal year month TEU_port_m share_i_p_q TEU_i_m
Pi_teu_per_hour_i_y H_annual_i_y w_i_m|y L_hours_i_m
```

- `labor_hours_QA.tsv`

```
port terminal year quarter sum_TEU_i_m_vs_TEU_i_q_ok annual_TEU_ok
annual_hours_ok preopen_zero_ok
```

- `labor_proxy_alt_calls_teams.tsv` (optional)

8) Reproducibility & Auditability

- Every `Pi_{i,y}` value is copied **verbatim** from published charts (screenshots archived).
- Terminal mapping (Hebrew → canonical) is included.
- Code stores exact TEU panel versions (monthly/quarterly) and their hashes.

9) Summary

This methodology converts **published annual productivity** and **observed throughput** into **terminal-monthly labor input** without smoothing across structural breaks, without interpolating shares within quarters, and with built-in QA and robustness paths. It is **transparent**, **replicable**, and **defensible** for event-study and mediation analyses.