Nonparametric Tests of Dynamic Complementarity (Extract)

James J. Heckman and Jin Zhou

Econ 350, Spring 2022

THE UNIVERSITY OF CHICAGO

Introduction

- Dynamic complementarity is a central idea in human development.
- It characterizes how early learning experiences affect subsequent learning and achievement.
- It is a component of the answer to the basic question of how easy it is for those who start behind to catch up with early starters, a fundamental problem in social policy.

- The technology of skill formation states that skill capital at age $a+1$, $K(a+1)$, depends on previous skill capital and investment:

$$K(a+1) = f^{(a)}(K(a), I(a)), \tag{1}$$

where $I(a)$ is investment.

- Knowledge is self productive $f_1^{(a)} > 0$ (skill begets skill).
- Investment is productive $f_2^{(a)} > 0$.

- Complementarity between skill capital and investment is found in many studies: $f_{12}^{(a)} > 0$ (those who know more learn more).
- Dynamic complementarity captures the notion that investment at age a (i.e., $I(a)$) raises the productivity of later investment (i.e., $I(a+j)$, $j > 0$):

$$\frac{\partial^2 K(a+j+1)}{\partial I(a) \partial I'(a+j)} > 0, \tag{2}$$

so early investment makes later life investment more productive.

Formal Tests of Dynamic Complementarity

- All inputs are assumed to have positive marginal products.
- Investment promotes the development of skills:

$$\frac{\partial f^{(a)}}{\partial I(a)} \Big|_{K(a)=\bar{K}} > 0.$$

- Its productivity can depend on the level of the capital stock $K(a)$.
- Age a^* is said to be sensitive for investment relative to a if investment is especially productive at a^*:

$$\underbrace{f_2^{(a)}(K(a), I(a)) \Big|_{a=a^*, K(a)=\bar{K}}}_{\text{Marginal productivity of investment at } a^*} > \underbrace{f_2^{(a)}(K(a), I(a)) \Big|_{a\neq a^*, K(a)=\bar{K}}}_{\text{Marginal productivity of investment at } a\neq a^*}. \tag{3}$$

- Dynamic Complementarity is defined assuming investment is productive (i.e., $f_2^{(a)}(K(a), I(a)) > 0$) and that there is complementarity between the stock of skills and investment (i.e., $f_{2,1}^{(a)}(K(a), I(a)) > 0$).
- Substituting recursively in (1):

$$\begin{aligned} K(a+j+1) &= f^{(a+j)}(K(a+j), I(a+j)) \\ &= f^{(a+j)}(f^{(a+j-1)}(K(a+j-1), I(a+j-1)), I(a+j)) \\ &= f^{(a+j)}(f^{(a+j-1)}(\ldots, f^{(a)}(K(a), I(a)), I(a+j)). \end{aligned}$$

- The marginal product of investment ($I(a)$) at age a on $K(a+j+1)$ depends on the sensitivity of investment at a:

$$\frac{\partial K(a+j+1)}{\partial I(a)} = \underbrace{\left[\prod_{\ell=1}^{j} f_1^{(a+\ell)}\right]}_{D(a+j+1)\text{: Transmission of } I(a) \text{ to } K(a+j+1)} \underbrace{f_2^{(a)}(K(a), I(a))}_{\substack{\text{Period } a \\ \text{Productivity: } P(a)}} > 0,$$

where $P(a)$ is the marginal productivity of investment and $D(a+j+1)$ is the transmission of period a investment to capital in period $a+j+1$.

- Dynamic complementarity arises from investment at age a on the productivity of future investments:

$$\frac{\partial^2 K(a+j+1)}{\partial I(a)\partial I'(a+j)} = \underbrace{f_{12}^{(a+j)}(K(a+j),I(a+j))}_{\text{Complementarity at age } a+j} \underbrace{\prod_{\ell=1}^{j-1} f_1^{(a+\ell)}}_{D(a+j)} \underbrace{f_2(K(a),I(a))}_{P(a)}$$

and

$$\frac{\partial K(a+j+1)}{\partial I(a)} = D(a+j+1)P(a).$$

- Dynamic complementarity depends on:
  - The productivity of investment in period a (i.e., $\boldsymbol{P(a)}$)
  - The appreciation or depreciation of period $\boldsymbol{I(a)}$ investment on capital (i.e., $\boldsymbol{D(a+j)}$)
  - The complementarity between capital and investment in period $a+j$ (i.e., $\boldsymbol{f_{12}^{(a+j)}(K(a+j), I(a+j))}$)

- Consider agents who start at a relatively insensitive age for productivity, $\alpha' > \alpha$, so that $\boldsymbol{P(\alpha')} < \boldsymbol{P(\alpha)}$.
- Everything else equal, later starting people who miss sensitive period α, find it hard to catch up.
- If transmission factor $\boldsymbol{D(\alpha + j)}$ is such that $\boldsymbol{D(\alpha')} < \boldsymbol{D(\alpha)}$, it could explain a lack of catch-up.
- In this case, late-starting children would lack preconditions for effective development.
- Lack of catch-up is evidence of dynamic complementarity.

- In the China REACH program, children enter the program at different ages.
- The curriculum is the same at each age, irrespective of the entry age of the children.
- For example, if child A enters at 6 months old, he will start with tasks designed for 6-months-old children.
- If child B enrolls at 10 months old, he will start with tasks designed for 10-months-old children, even though he has not been exposed to the earlier parts of the program.
- In other words, child B misses the tasks for 6- to 10-month-old children.
- This setting provides us the ideal environment to test dynamic complementarity of investment.

We Formalize This Idea as Follows:

- Let $a_j^*(s, \ell)$ be the age of entry into the program for agent j at level ℓ for skill s.
- $a_k^*(s, \ell)$ is the age of entry into the program for agent k.
- $Q_j(a(s, \ell))$ and $Q_k(a'(s, \ell))$, for $a \neq a'$, are passing rates on a test of skill for the two agents.

- Compute $E(Q_j(a(s,\ell)) \mid a(s,\ell), a(s,\ell) > a_j^*(s,\ell))$ and $E(Q_k(a(s,\ell)) \mid a(s,\ell), a(s,\ell) > a_k^*(s,\ell))$.
- For $a_k^*(s,\ell) > a_j^*(s,\ell)$, dynamic complementarity implies

$$\begin{aligned} \boldsymbol{H_0}: \; & E(Q_k(a(s,\ell)) \mid a(s,\ell), a(s,\ell) \geq a_k^*(s,\ell)) \\ & < E(Q_j(a(s,\ell)) \mid a(s,\ell), a(s,\ell) \geq a_j^*(s,\ell)), \quad \text{all } (s,\ell). \end{aligned}$$

- To test dynamic complementarity, we first categorize three groups based on the children's ages at enrollment:
  1. Ages 10–15 months
  2. Ages 16–20 months
  3. Ages 21–25 months
- Then, we compare children's passing rate across the three groups.

Tests Based on Each Task Item

Table 1: Passing Rates on Language Tasks by Enrollment Age Group and Ability

	Task Index	1	2	3	4	5	6	7	8	9	10	11	12	13
Fast	Age 10–15	0.815	0.654	0.741	0.893	0.857	1.000	0.958	0.833	0.833	0.958	0.840	0.875	0.926
	Age 16–20					0.882	1.000	0.941	0.900	0.947	0.950	0.947	0.947	0.885
	p-value: (10–15)> (16–20)					0.594		0.407	0.738	0.884	0.449	0.876	0.796	0.308
Normal	Age 10-15	0.333	0.396	0.324	0.624	0.505	0.646	0.510	0.483	0.548	0.754	0.491	0.649	0.685
	Age 16–20	0.250	0.250	0.188	0.529	0.295	0.522	0.289	0.333	0.352	0.672	0.397	0.500	0.455
	p-value: (10–15) > (16–20)	0.274	0.115	0.108	0.239	**0.007**	**0.078**	**0.007**	**0.026**	**0.008**	0.130	0.122	**0.038**	**0.001**
Slow	Age 10–15	0.065	0.133	0.114	0.152	0.049	0.233	0.130	0.319	0.083	0.348	0.133	0.156	0.255
	Age 16–20					0.000	0.067	0.000	0.053	0.000	0.222	0.188	0.133	0.150
	p-value: (10–15) > (16–20)					**0.079**	**0.040**	**0.006**	**0.002**	**0.021**	0.156	0.683	0.417	0.158

- Table 1 compares the passing rate for each language task item.
- Because the children were evaluated on the tasks in Table 1 before they were 20 months old, there is no data for the 21–25 age group.
- We can still compare children in the 10–15 and 16–20 age groups for these tasks.
- We find that dynamic complementarity effects are stronger for children in the normal and slow groups.
- But children in the fast group do not perform significantly differently on any task irrespective of the age of entry.

- When we compare fast children in the 10–15 and 16–20 age groups, we find that 29% of tasks demonstrate significant dynamic complementarity effects.
- The numbers for normal and slow children are 40% and 44%, respectively.
- When we compare fast children in the 10–15 and 21–25 groups, we find that 25% of tasks demonstrate significant dynamic complementarity effects.

- Dynamic complementarity may not be universal across skills.
- The numbers for normal and slow children are 44% and 57%, respectively.
- In general, we find that children in the normal and slow groups exhibit more statistically significant dynamic complementarity effects.
- We also plot the passing rate for each task item by ability group.

Figure 1: Passing Rate for Language Tasks for **Fast Group** by Enrollment Age

(a) Age 10–15 vs. Age 16–20

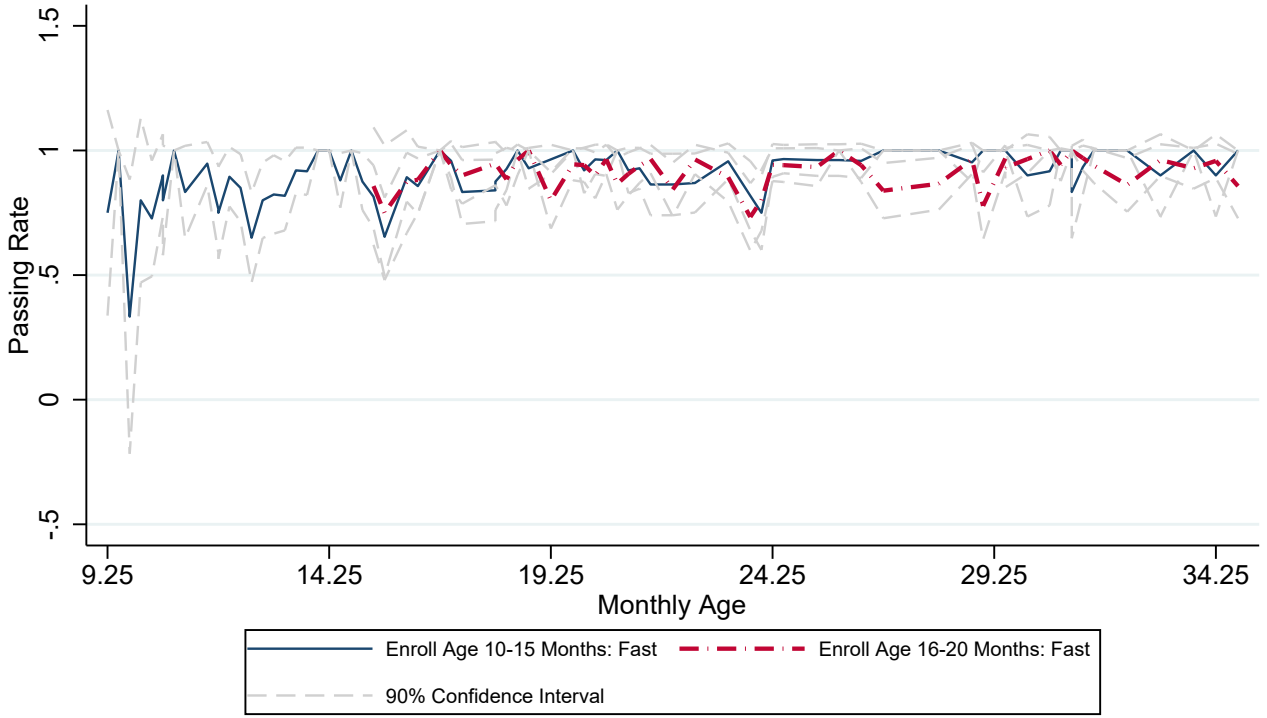


Figure 1: Passing Rate for Language Tasks for **Fast Group** by Enrollment Age, Cont'd

(b) Age 10–15 vs. Age 21–25

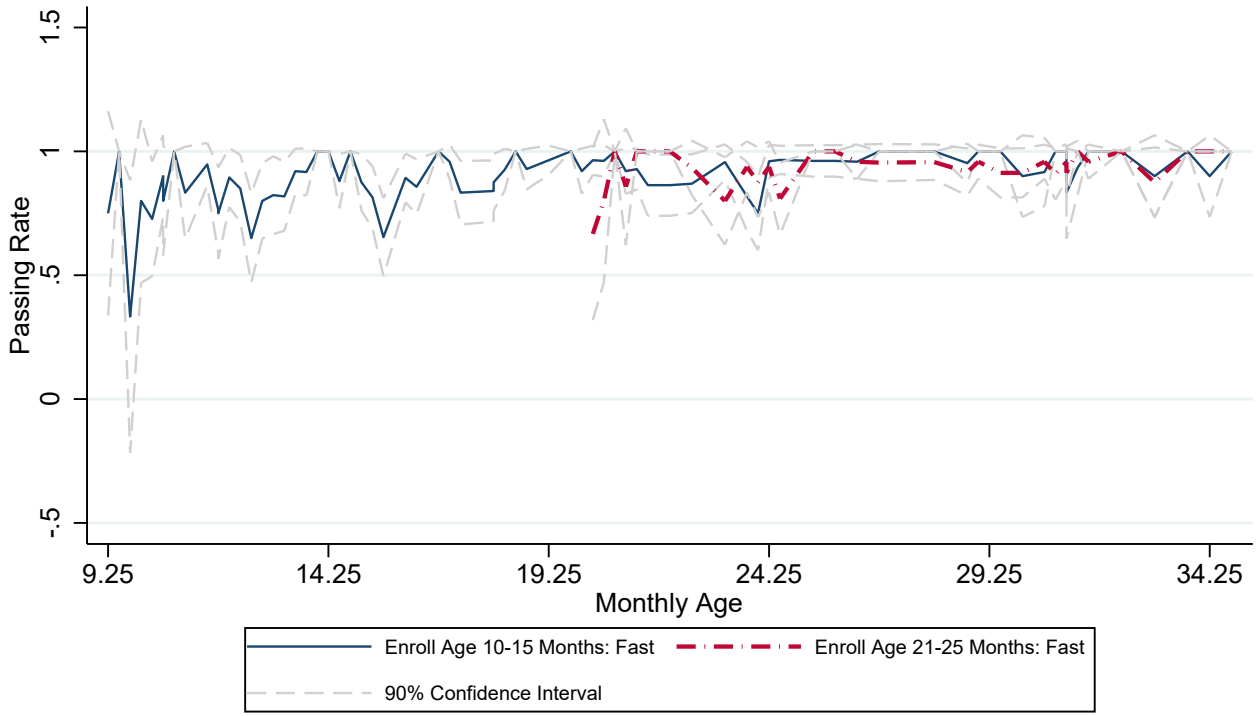


Figure 1: Passing Rate for Language Tasks for **Fast Group** by Enrollment Age, Cont'd

(c) Age 16–20 vs. Age 21–25

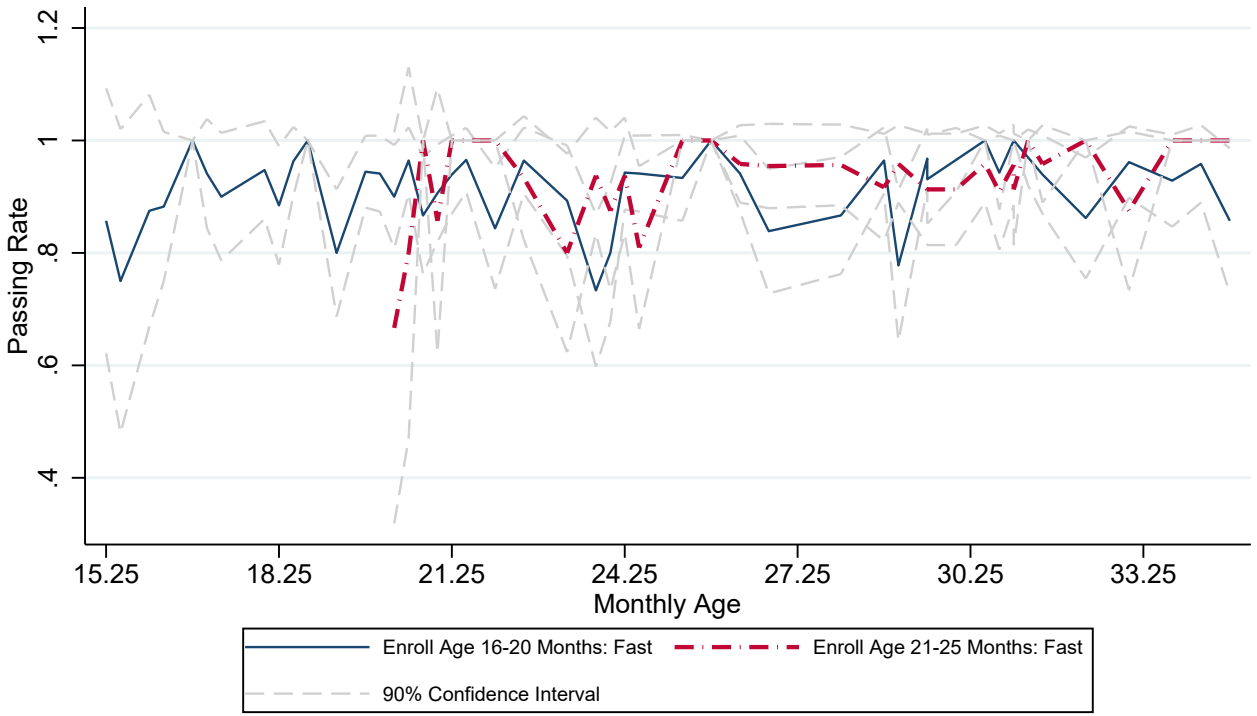


Figure 2: Passing Rate for Language Tasks for **Normal Group** by Enrollment Age

(a) Age 10–15 vs. Age 16–20

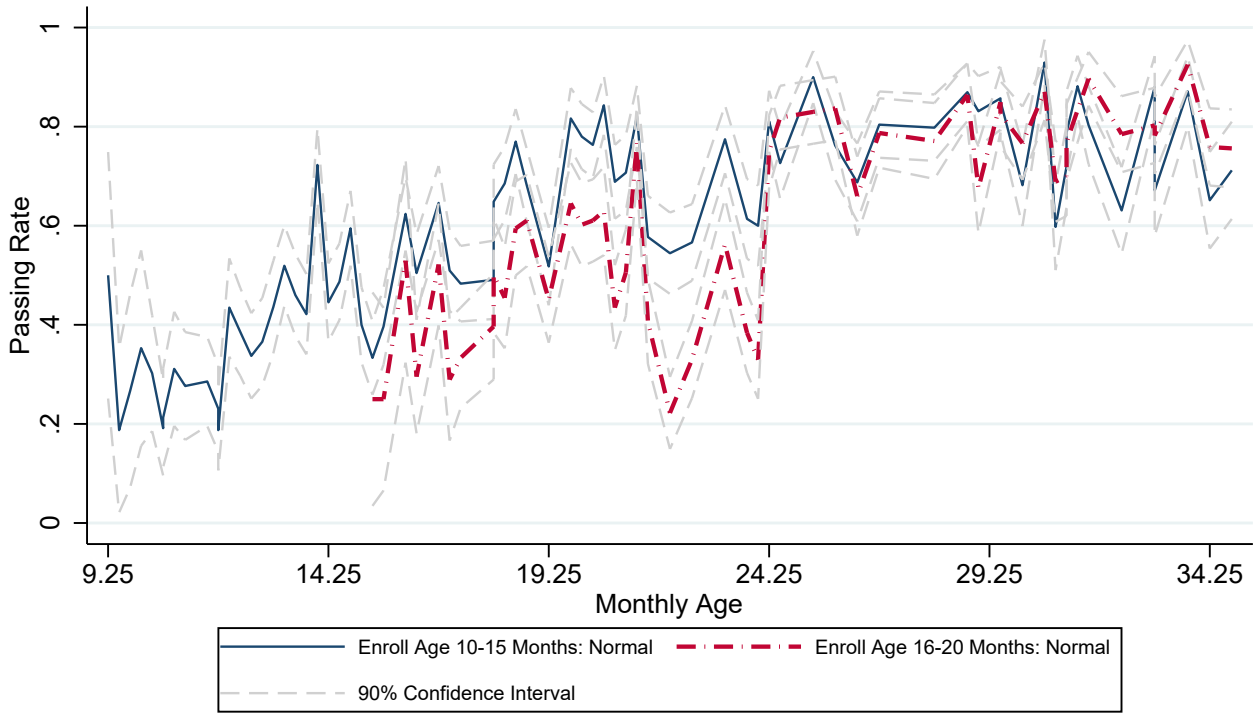


Figure 2: Passing Rate for Language Tasks for **Normal Group** by Enrollment Age, Cont'd

(b) Age 10–15 vs. Age 21–25

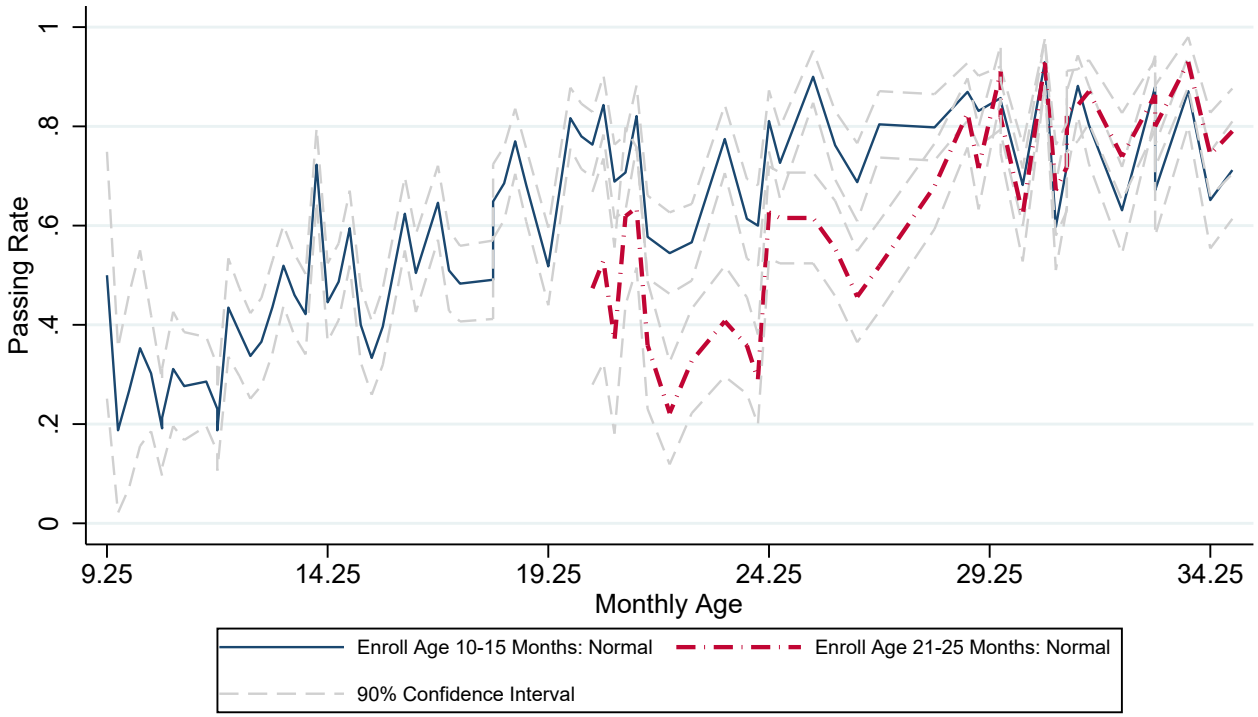


Figure 2: Passing Rate for Language Tasks for **Normal Group** by Enrollment Age, Cont'd

(c) Age 16–20 vs. Age 21–25

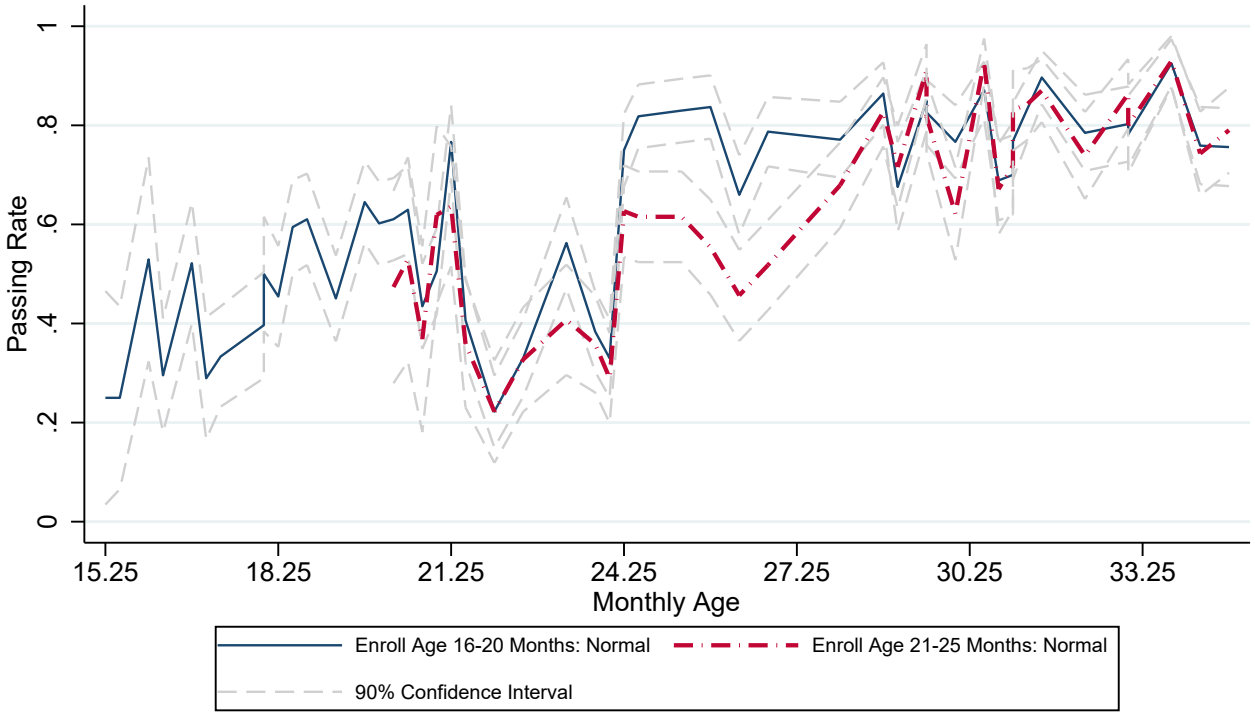


Figure 3: Passing Rate for Language Tasks for **Slow Group** by Enrollment Age

(a) Age 10–15 vs. Age 16–20

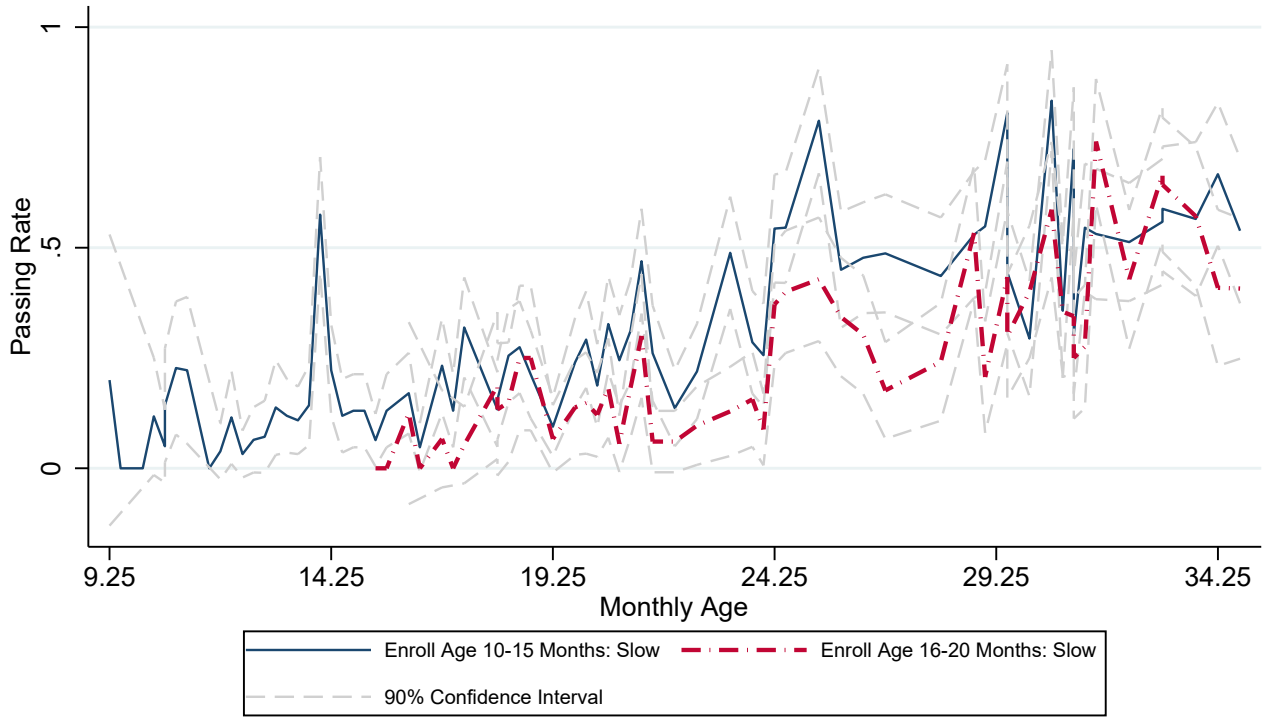


Figure 3: Passing Rate for Language Tasks for **Slow Group** by Enrollment Age, Cont'd

(b)Age 10–15 vs. Age 21–25

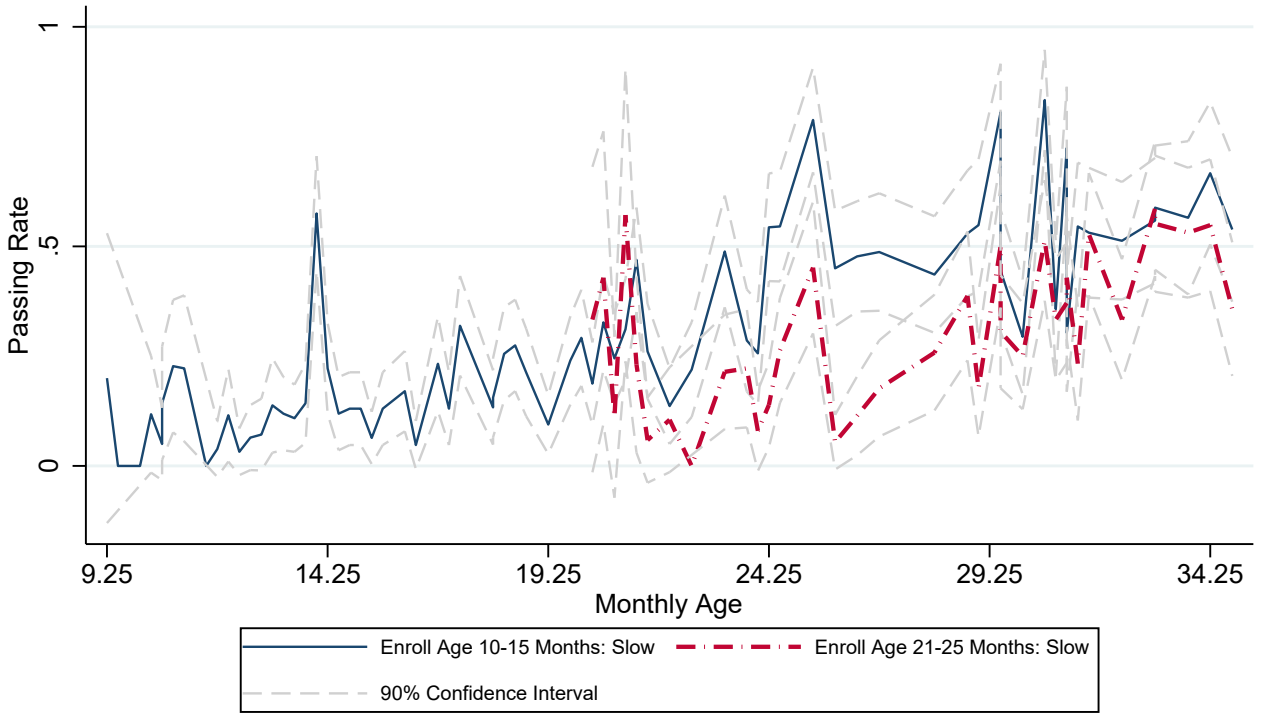
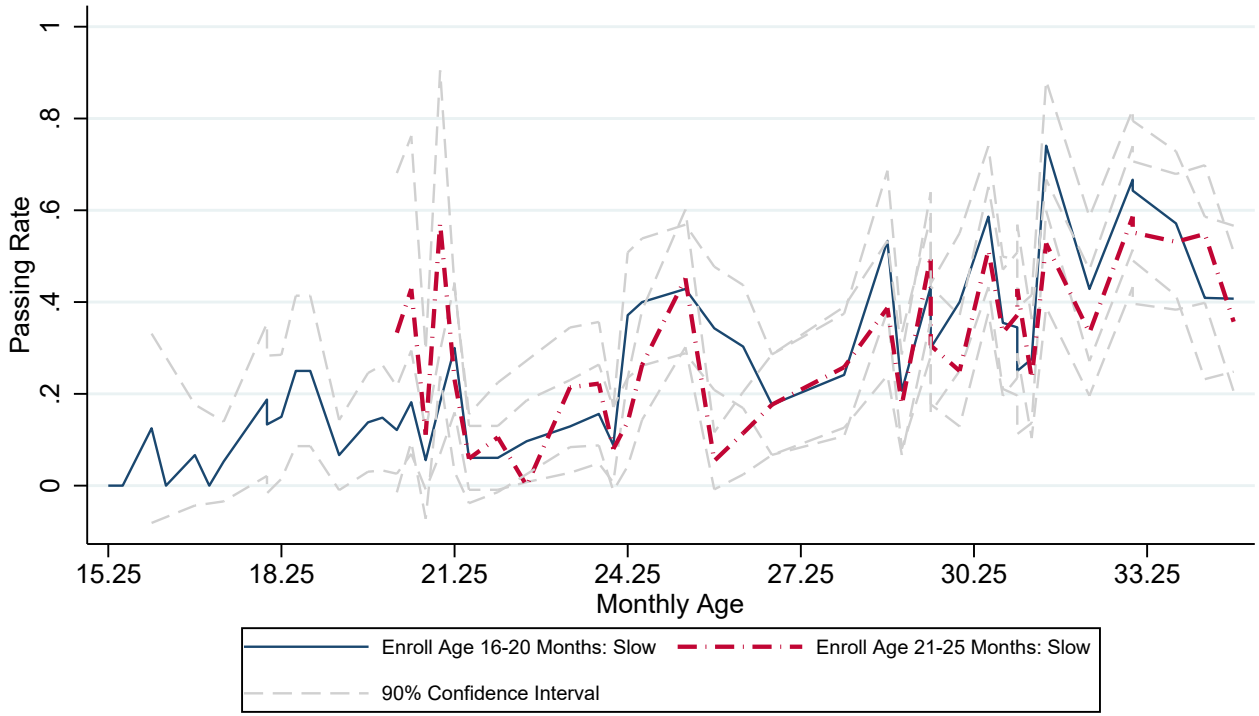


Figure 3: Passing Rate for Language Tasks for Slow Group by Enrollment Age, Cont'd

(a) Age 16–20 vs. Age 21–25



Tests Based on the Measures by Difficulty Level

- We test for dynamic complementarity based on the measures across difficulty levels.
- For each difficulty level, multiple tasks start at different weekly ages.

Table 2: Monthly Age Ranges in the Curriculum Design

Difficulty Level	First Task	Last Task
2	6.75	20.00
3	9.50	18.25
4	10.00	18.50
5	10.50	15.50
6	10.75	25.25
7	19.25	31.50
8	21.75	40.75
9	26.00	42.75
10	26.00	39.00
11	34.00	42.50

- Table 2 documents the monthly ages for the first and last tasks at each level.
- Because we group the children by their age at enrollment (i.e., 10–15 months, 16–20 months, and 21–25 months), the measures of the first six difficulty levels can be significantly affected by the age differences.
- The effects of dynamic complementarity should be examined using measures evaluated when children have the same age but with different lengths of program exposure.
- Therefore, the higher difficulty level measures (e.g., from difficulty level 7) are more suitable for us to accurately examine dynamic complementarity effects.

Tests Based on Average Passing Rate Performance

Table 3: Language Passing Rate by Enrollment Age

	Language Level				
	7	8	9	10	11
	Enroll (10–15) vs. (16–20)				
Age 10–15	0.652	0.715	0.726	0.739	0.728
Age 16–20	0.629	0.693	0.702	0.765	0.756
p-value	0.376	0.344	0.502	0.515	0.359
N	427	424	382	302	404
	Enroll (10–15) vs. (21–25)				
Age 10–15	0.652	0.715	0.726	0.739	0.728
Age 21–25	0.627	0.694	0.637	0.813	0.649
p-value	0.365	0.393	**0.022**	**0.041**	**0.017**
N	382	382	348	298	364
	Enroll (16–20) vs. (21–25)				
Age 16–20	0.629	0.693	0.702	0.765	0.756
Age 21–25	0.627	0.694	0.637	0.813	0.649
p-value	0.928	0.957	0.072	0.133	**0.001**
N	401	404	374	332	390

- Table 3 presents the test of the mean passing rate of language skill at each difficulty level for different enrollment groups.
- We do not find strong dynamic complementarity effects when comparing the groups who enroll in the program between the ages of 10 and 15 months vs. those who enroll between the ages of 16 and 20 months.
- The comparison between the 10–15 group and the 21–25 one is slightly stronger.
- For example, we find significant dynamic complementarity effects at difficulty levels 9 and 11.
- Similarly, we find significant effects for the comparison between the 16–20 and 21–25 groups.

- Children perform completely differently by ability group.
- For example, children in the fast group perform persistently well regardless of enrollment age.
- To eliminate the impacts across different ability groups, we conduct a similar exercise by ability group.

Table 4: Language Passing Rate by Enrollment Age and Ability

	Language Level				
	7	8	9	10	11
			Fast		
			Enroll (10–15) vs. (16–20)		
Age 10–15	0.932	0.896	0.947	0.950	0.944
Age 16–20	0.892	0.919	0.897	0.911	0.979
p-value	0.168	0.592	0.250	0.569	0.235
N	67	67	58	41	63
			Enroll (10–15) vs. (21–25)		
Age 10–15	0.932	0.896	0.947	0.950	0.944
Age 21–25	0.936	0.935	0.949	0.938	0.922
p-value	0.892	0.396	0.951	0.769	0.567
N	53	56	50	41	52
			Enroll (16–20) vs. (21–25)		
Age 16–20	0.892	0.919	0.897	0.911	0.979
Age 21–25	0.936	0.935	0.949	0.938	0.922
p-value	0.151	0.587	0.190	0.596	**0.028**
N	68	71	64	54	67

Table 4: Language Passing Rate by Enrollment Age and Ability, Cont'd

	Language Level				
	7	8	9	10	11
			Normal		
			Enroll (10–15) vs. (16–20)		
Age 10–15	0.702	0.739	0.746	0.755	0.767
Age 16–20	0.632	0.674	0.740	0.806	0.785
p-value	**0.001**	**0.011**	0.902	0.307	0.613
N	230	229	208	172	220
			Enroll (10–15) vs. (21–25)		
Age 10–15	0.702	0.739	0.746	0.755	0.767
Age 21–25	0.651	0.726	0.635	0.852	0.692
p-value	**0.036**	0.583	**0.026**	**0.033**	**0.046**
N	202	201	184	162	194
			Enroll (16–20) vs. (21–25)		
Age 16–20	0.632	0.674	0.740	0.806	0.785
Age 21–25	0.651	0.726	0.635	0.852	0.692
p-value	0.437	**0.039**	**0.024**	0.210	**0.007**
N	210	210	198	180	206

Table 4: Language Passing Rate by Enrollment Age and Ability, Cont'd

	Language Level				
	7	8	9	10	11
			Slow		
			Enroll (10–15) vs. (16–20)		
Age 10–15	0.336	0.508	0.495	0.566	0.455
Age 16–20	0.233	0.394	0.332	0.414	0.370
p-value	**0.017**	**0.019**	**0.050**	0.117	0.206
N	93	91	83	62	85
			Enroll (10–15) vs. (21–25)		
Age 10–15	0.336	0.508	0.495	0.566	0.455
Age 21–25	0.290	0.376	0.320	0.556	0.253
p-value	0.301	**0.007**	**0.036**	0.907	**0.001**
N	94	92	83	70	85
			Enroll (16–20) vs. (21–25)		
Age 16–20	0.233	0.394	0.332	0.414	0.370
Age 21–25	0.290	0.376	0.320	0.556	0.253
p-value	0.139	0.733	0.863	**0.091**	**0.066**
N	83	83	76	62	78

- Table 4 shows the test results by ability group.
- We find that for children in the fast group, there is no evidence of dynamic complementarity across difficulty levels.
- However, we find stronger dynamic complementarity effects for normal and slow groups.
- For example, we find significant effects at levels 7 and 8 when comparing the normal children in the 10–15 and 16–20 month age groups.
- We find similarly significant effects at levels 7, 8, and 9 for slow children.

- Also, when comparing slow children in the 10–15 and 21–25 age groups, there are significant impacts at levels 8, 9, and 11.
- We find similar results when comparing the 10–15 and 21–25 age groups, and the 16–20 and 21–25 age groups.
- These results indicate strong dynamic complementarity effects for children not in the fast group.

Tests Based on Time to First Mastery

- The average passing rate is a traditional measure of child learning and knowledge, but other measures also capture aspects of a child's knowledge and comprehension ability.
- We also examine evidence on dynamic complementarity using time to first mastery as a measure of knowledge and learning.
- This measure captures how rapidly the child can master tasks.
- Time to first mastery is the number of trials a child takes until the first success at each difficulty level during the intervention for each skill type.

Table 5: Time to First Mastery of Language Tasks by Enrollment Age

	Language Level				
	7	8	9	10	11
	Enroll (10–15) vs. (16–20)				
Age 10–15	2.380	2.030	1.355	1.212	1.286
Age 16–20	2.758	2.316	1.399	1.278	1.278
p-value	0.138	0.058	0.640	0.434	0.910
N	419	412	338	264	380
	Enroll (10–15) vs. (21–25)				
Age 10–15	2.380	2.030	1.355	1.212	1.286
Age 21–25	2.213	2.379	1.690	1.396	1.702
p-value	0.420	**0.027**	**0.001**	**0.040**	**0.000**
N	374	374	300	272	336
	Enroll (16–20) vs. (21–25)				
Age 16–20	2.758	2.316	1.399	1.278	1.278
Age 21–25	2.213	2.379	1.690	1.396	1.702
p-value	**0.023**	0.715	**0.003**	0.201	**0.000**
N	393	392	328	310	366

- Table 5 reports the mean of time to first mastery at each difficulty level for language skill for different enrollment age groups (measured in months).
- The smaller values correspond to more rapid learning.
- We find significant dynamic complementarity effects at levels 8–11 when we compare the 10–15 and 21–25 age groups.
- We also find significant effects at levels 9 and 11 when comparing the 16–20 and 21–25 groups.

Table 6: Language Time to First Mastery by Enrollment Age and Ability

	Language Level				
	7	8	9	10	11
		Fast			
		Enroll (10–15) vs. (16–20)			
Age 10–15	1.038	1.160	1.045	1.143	1.083
Age 16–20	1.220	1.171	1.056	1.000	1.000
p-value	0.097	0.926	0.869	0.050	0.069
N	67	66	58	40	63
		Enroll (10–15) vs. (21–25)			
Age 10–15	1.038	1.160	1.045	1.143	1.083
Age 21–25	1.148	1.100	1.036	1.037	1.036
p-value	0.179	0.572	0.865	0.227	0.473
N	53	55	50	41	52
		Enroll (16–20) vs. (21–25)			
Age 16–20	1.220	1.171	1.056	1.000	1.000
Age 21–25	1.148	1.100	1.036	1.037	1.036
p-value	0.540	0.453	0.715	0.331	0.241
N	68	71	64	53	67

Table 6: Language Time to First Mastery by Enrollment Age and Ability, Cont'd

	Language Level				
	7	8	9	10	11
			Normal		
			Enroll (10–15) vs. (16–20)		
Age 10–15	1.820	1.824	1.438	1.277	1.232
Age 16–20	2.378	2.362	1.433	1.227	1.183
p-value	**0.014**	**0.000**	0.969	0.600	0.468
N	230	224	193	153	214
			Enroll (10–15) vs. (21–25)		
Age 10–15	1.820	1.824	1.438	1.277	1.232
Age 21–25	2.198	2.278	1.763	1.381	1.489
p-value	0.055	**0.003**	**0.032**	0.359	**0.003**
N	202	198	165	149	187
			Enroll (16–20) vs. (21–25)		
Age 16–20	2.378	2.362	1.433	1.227	1.183
Age 21–25	2.198	2.278	1.763	1.381	1.489
p-value	0.462	0.599	**0.022**	0.127	**0.000**
N	210	206	180	172	203

Table 6: Language Time to First Mastery by Enrollment Age and Ability, Cont'd

	Language Level				
	7	8	9	10	11
			Slow		
			Enroll (10–15) vs. (16–20)		
Age 10–15	4.750	3.163	1.500	1.154	1.676
Age 16–20	6.432	4.028	1.880	2.100	2.033
p-value	**0.039**	0.076	0.147	**0.004**	0.180
N	85	85	55	46	67
			Enroll (10–15) vs. (21–25)		
Age 10–15	4.750	3.163	1.500	1.154	1.676
Age 21–25	3.211	3.974	2.417	1.871	3.407
p-value	**0.012**	0.064	**0.000**	**0.009**	**0.000**
N	86	88	54	57	64
			Enroll (16–20) vs. (21–25)		
Age 16–20	6.432	4.028	1.880	2.100	2.033
Age 21–25	3.211	3.974	2.417	1.871	3.407
p-value	**0.000**	0.927	0.065	0.573	**0.001**
N	75	75	49	51	57

- We also examine the dynamic complementarity effects by ability group in Table 6
- We find results similar to those of the average passing rate.
- Children in the normal and slow groups show more evidence of dynamic complementarity, especially when comparing children in the 10–15 and 21–25 age groups, since there may be a ten-month difference in intervention exposure.
- But for children in the fast group, we do not find any statistically significant dynamic complementarity effects across different enrollment age groups.

Tests Based on Denver Tests

- The preceding sections presented tests based on the UHP weekly measures.
- We use two different ways to examine the dynamic complementarity effects using the Denver tests.
- The difference between the Denver test performance outcomes and UHP task outcomes is due to evaluation frequency.

- For the Denver test, the children are only evaluated at the midline (after about 9–10 months of intervention exposure) and at the endline (after about 21–22 months of intervention exposure).
- The UHP framework produces weekly measures for the children in the treatment group.
- Therefore, when we use the Denver test results, we need to compare the children of the same ages but with different intervention exposure times.

- We focus our samples on monthly ages between 29 and 39 months to conduct an analysis of dynamic complementarity using raw Denver scores.
- The samples include four components: control group children within this age range at the midline, control group children within this age range at the endline, treatment group children within this age range at the midline, and treatment group children within this age range at the endline.
- Since control group children do not receive any intervention, we can pool the two groups of control children together as a comparison group for this age range.
- First, we evaluate the treatment effects for these children.

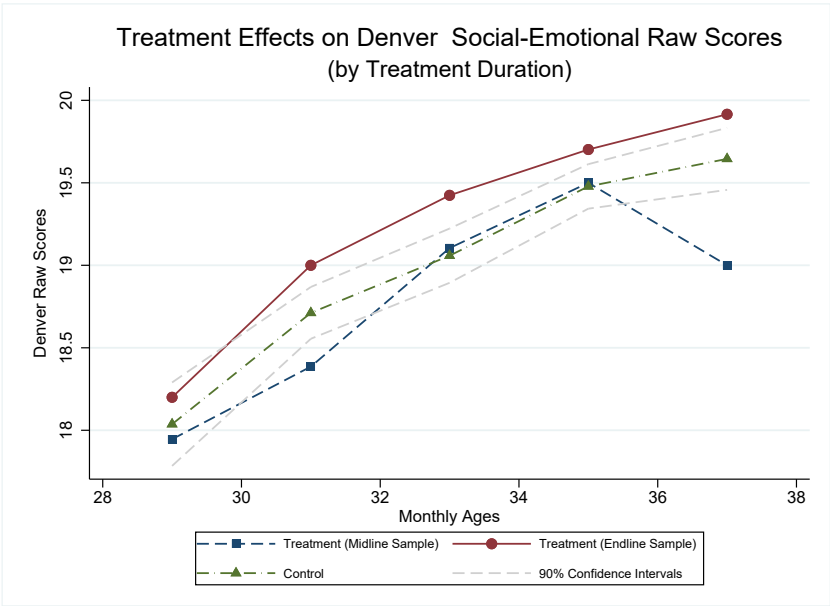
Table 7: Treatment Effects on Denver Raw Scores (Monthly Age Range: 29–39)

	Social-Emotional	Fine Motor	Language and Cognitive	Gross Motor
		Midline Treatment vs. Control		
Treatment	-0.056	0.578	0.683**	0.081
	[-0.248, 0.125]	[-0.117, 1.442]	[0.194, 1.245]	[-0.347, 0.438]
		Endline Treatment vs. Control		
Treatment	0.260**	0.177	0.782**	0.024
	[0.084, 0.439]	[-0.075, 0.436]	[0.261, 1.293]	[-0.214, 0.261]

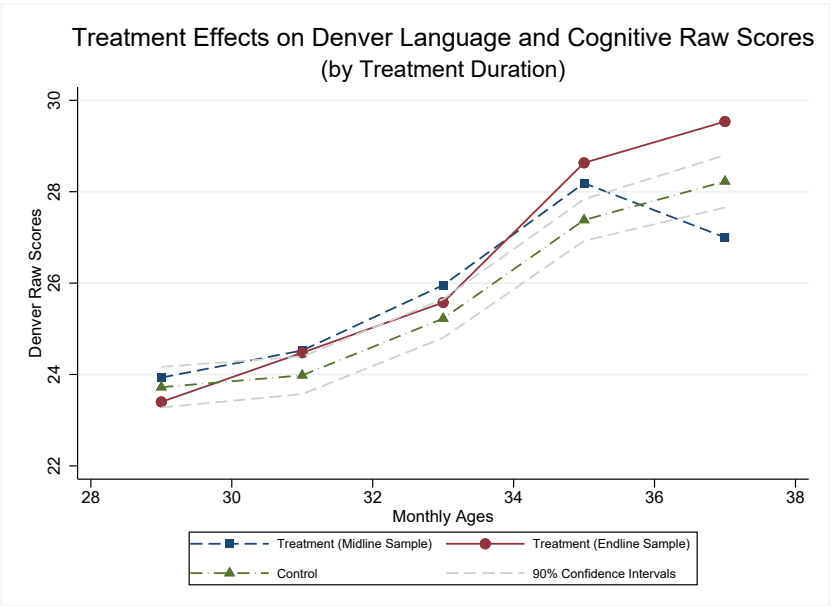
- Table 7 reports the treatment effects on Denver raw scores for the two treatment groups with different intervention exposure time.
- The group with longer treatment duration shows more significant treatment effects for language and cognitive scores and social-emotional scores.
- These results also support the existence of dynamic complementarity effects.
- Next, we consider treatment effects for each two-month bin age range by comparing the two treatment groups with different exposure lengths to the control group with the same monthly age range.

Figure 4: Treatment Effects on Denver Raw Scores by Monthly Age

(a) Social-Emotional



(b) Language and Cognitive



- Figure 4 plots the Denver raw scores for three groups: control group, treatment midline samples, and treatment endline samples.
- For each point, we use samples with the same monthly ages.
- For example, the estimates at 29 months include children whose ages are between 29 and 31 months.
- Our approach is similar for other age intervals.
- Estimates by each age range also support our dynamic complementarity effect results.

- Children with longer intervention exposure have better performance than those with shorter intervention exposure.
- We analyze the Denver raw scores.
- We conduct a robustness analysis based on the estimates of individual latent skills reported in Zhou, Heckman, Liu, and Lu (2021).
- These estimates eliminate fluctuations in the estimated scores and adjust for the task difficulty of each item.
- We can thus directly compare these latent skill scores across age enrollment groups.

Table 8: Latent Denver Scores (Midline) by Enrollment Age

	Language and Cognitive	Social-Emotional	Fine Motor	Gross Motor
		Enroll (10–15) vs. (16–20)		
Age 10–15	1.639	-1.300	1.889	1.805
Age 16–20	1.609	-2.358	0.379	-0.028
p-value	0.781	**0.000**	**0.000**	**0.000**
N	431	431	431	431
		Enroll (10–15) vs. (21–25)		
Age 10–15	1.639	-1.300	1.889	1.805
Age 21–25	0.658	-2.271	0.473	-0.813
p-value	**0.000**	**0.000**	**0.000**	**0.000**
N	427	427	427	427
		Enroll (16–20) vs. (21–25)		
Age 16–20	1.609	-2.358	0.379	-0.028
Age 21–25	0.658	-2.271	0.473	-0.813
p-value	**0.000**	0.671	0.542	**0.000**
N	388	388	388	388

Table 9: Latent Denver Scores (Endline) by Enrollment Age

	Language and Cognitive	Social-Emotional	Fine Motor	Gross Motor
		Enroll (10–15) vs. (16–20)		
Age 10–15	0.363	-3.013	0.375	-0.814
Age 16–20	-1.576	-4.024	0.254	0.043
p-value	**0.000**	**0.000**	0.370	**0.000**
N	368	368	368	368
		Enroll (10–15) vs. (21–25)		
Age 10–15	0.363	-3.013	0.375	-0.814
Age 21–25	-2.518	-3.881	0.018	0.852
p-value	**0.000**	**0.000**	**0.004**	**0.000**
N	368	368	368	368
		Enroll (16–20) vs. (21–25)		
Age 16–20	-1.576	-4.024	0.254	0.043
Age 21–25	-2.518	-3.881	0.018	0.852
p-value	**0.000**	0.277	0.059	**0.000**
N	332	332	332	332

- Tables 8–9 report the comparison of the difficulty-adjusted, measurement error–adjusted latent Denver scores across enrollment age groups estimated in Zhou, Heckman, Liu, and Lu (2021).
- These estimates correct for measurement error, adjust for item difficulty, and smooth out item and level fluctuations.
- We find a generally consistent pattern: children who enroll in the program at younger ages have higher latent Denver scores.
- Also, the earlier the children enroll in the program, the better their performance.
- The finding that the children who enroll in the program at younger ages have better performance than those who enroll at older ages supports dynamic complementarity.

Growth and Treatment Exposure Duration

Table 10: The Effects of Treatment Exposure Duration on Denver Raw Scores

	Social-Emotional	Fine Motor	Language and Cognitive	Gross Motor
		Midline Denver Scores		
Treatment Duration	0.013*	0.006	0.039*	-0.010
	[0.001, 0.024]	[-0.008, 0.020]	[0.005, 0.079]	[-0.032, 0.009]
		Endline Denver Scores		
Treatment Duration	-0.004	0.007*	0.042***	0.002
	[-0.009, 0.002]	[0.001, 0.016]	[0.028, 0.057]	[-0.007, 0.011]

- Table 10 reports the effects of treatment exposure duration on raw Denver scores.
- Longer treatment duration significantly improves Denver language and cognitive scores at both midline and endline, social-emotional scores at midline, and fine motor scores at endline, but effects are much less pronounced for other groups.