

ИЗОБРАЖЕНИЕ ГРАФИКА ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ РЕГРЕССИОННОЙ МОДЕЛИ

```
. reg EARNINGS S EXP
```

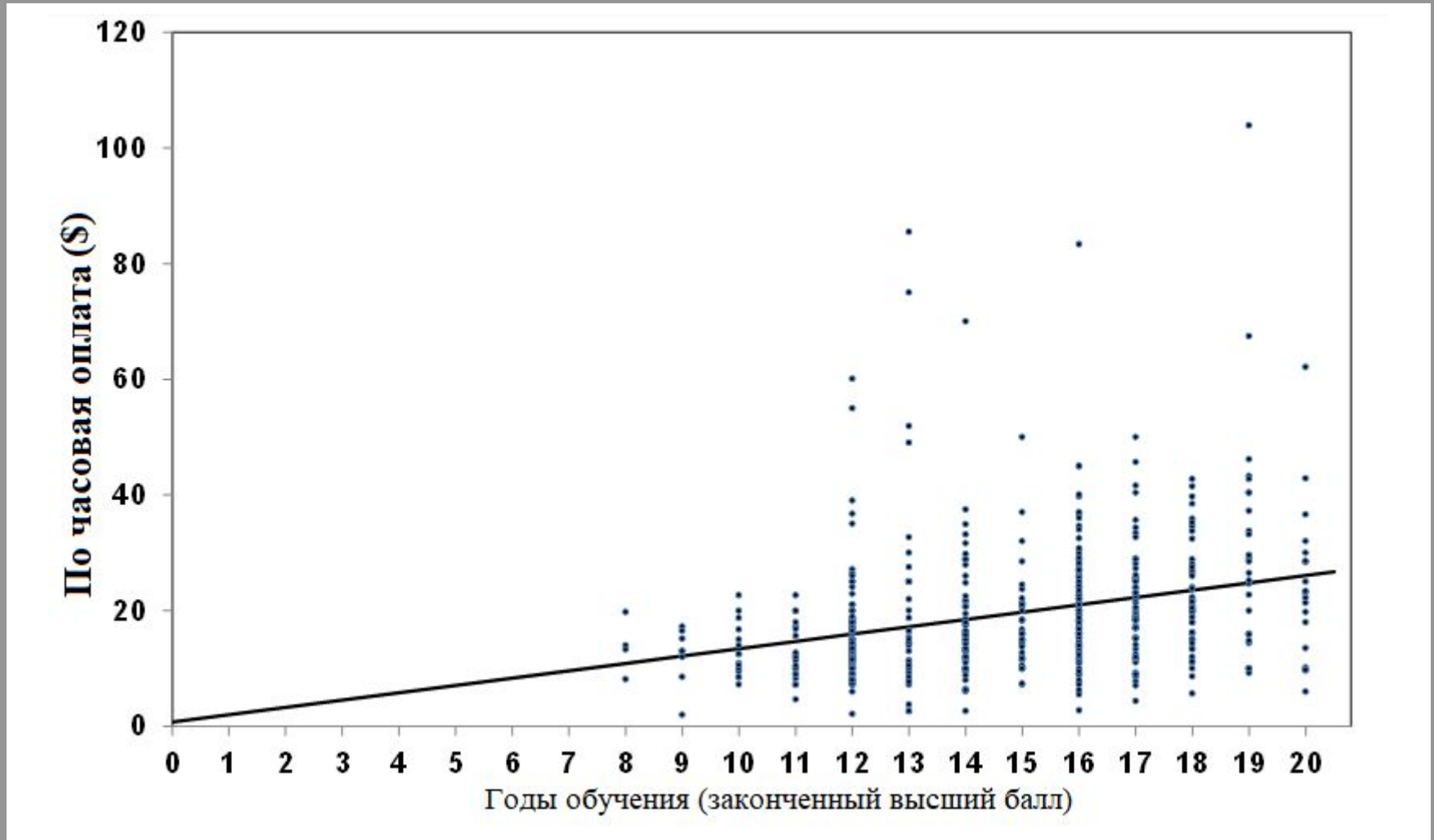
Начало	SS	df	MS	Number of obs = 500		
Модель	8735.42401	2	4367.712	F(2, 497)	=	35.24
Остаток	61593.5422	497	123.930668	Prob > F	=	0.0000
				R-squared	=	0.1242
				Adj R-squared	=	0.1207
Всего	70328.9662	499	140.939812	Root MSE	=	11.132

Доход	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
S	1.877563	.2237434	8.39	0.000	1.437964	2.317163
EXP	.9833436	.2098457	4.69	0.000	.5710495	1.395638
_cons	-14.66833	4.288375	-3.42	0.001	-23.09391	-6.242752

$$\text{ДОХОД} = -14.67 + 1.88S + 0.98EXP$$

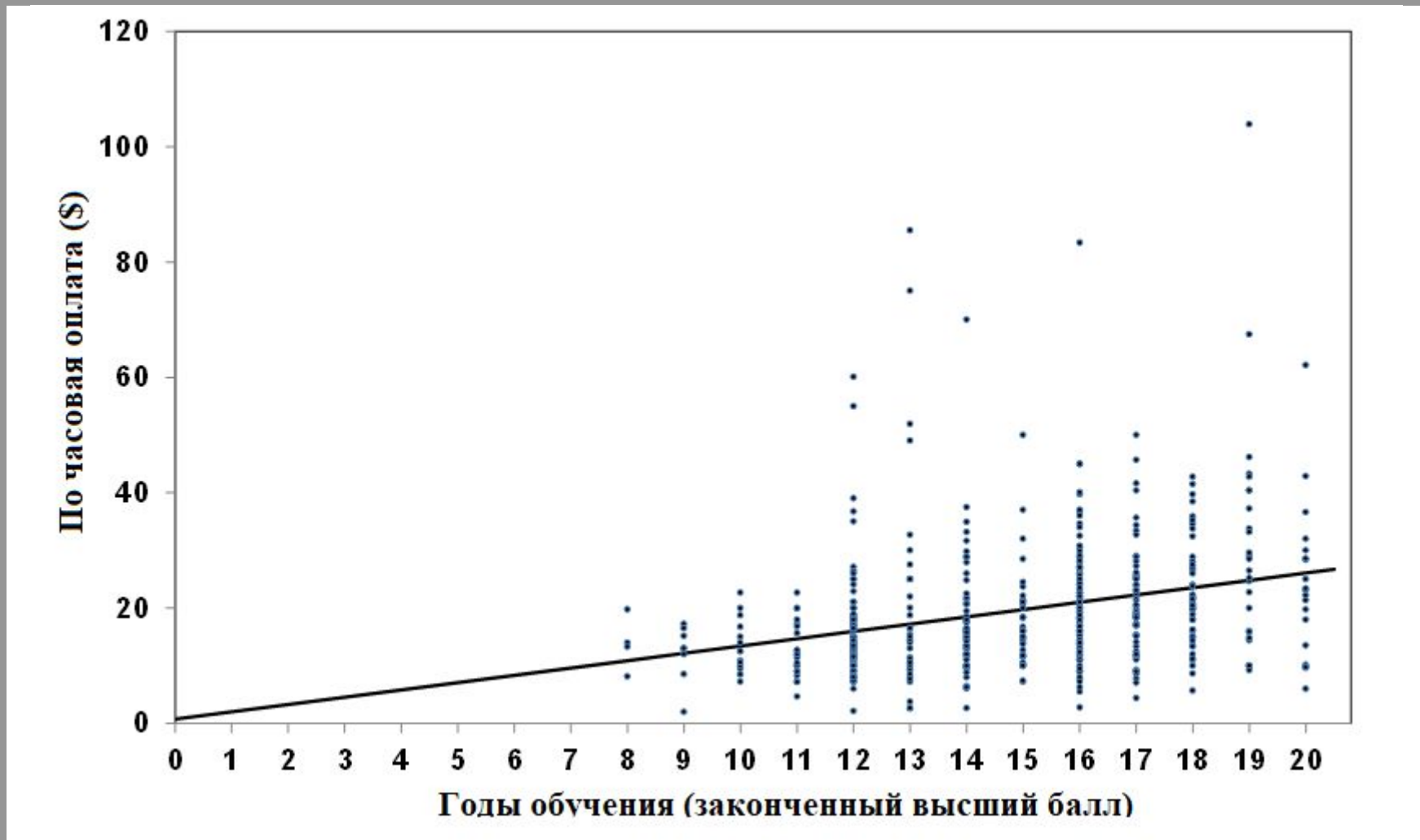
Продукция выше показывает результат регрессии ДОХОДА, почасового дохода в долларах умноженного на S, годы обучения и EXP, годы опыта работы.

ИЗОБРАЖЕНИЕ ГРАФИКА ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА



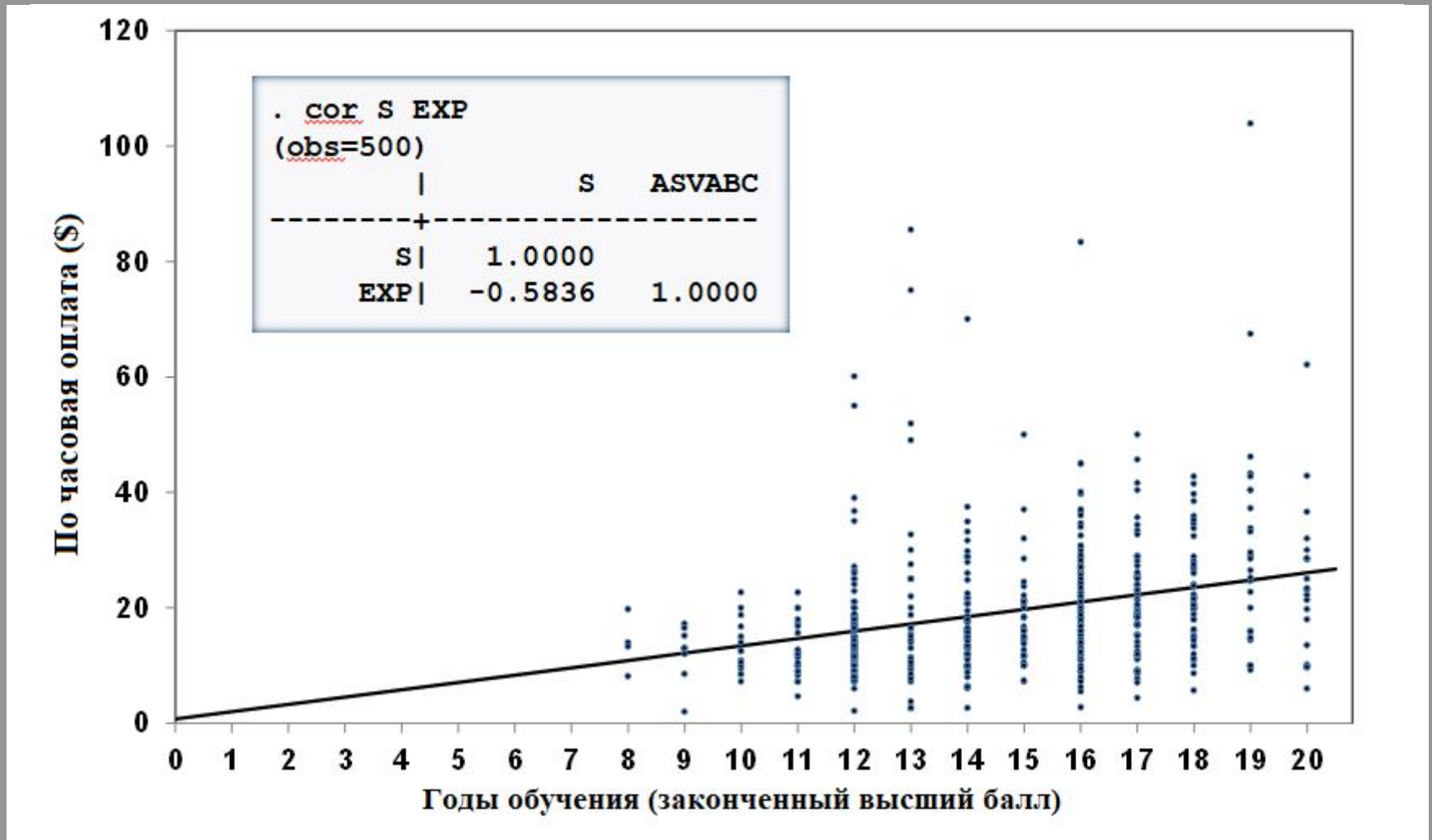
Предположим, что Вы интересовались отношением между ДОХОДОМ и S и хотели представить его графически, используя типовые данные.

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА



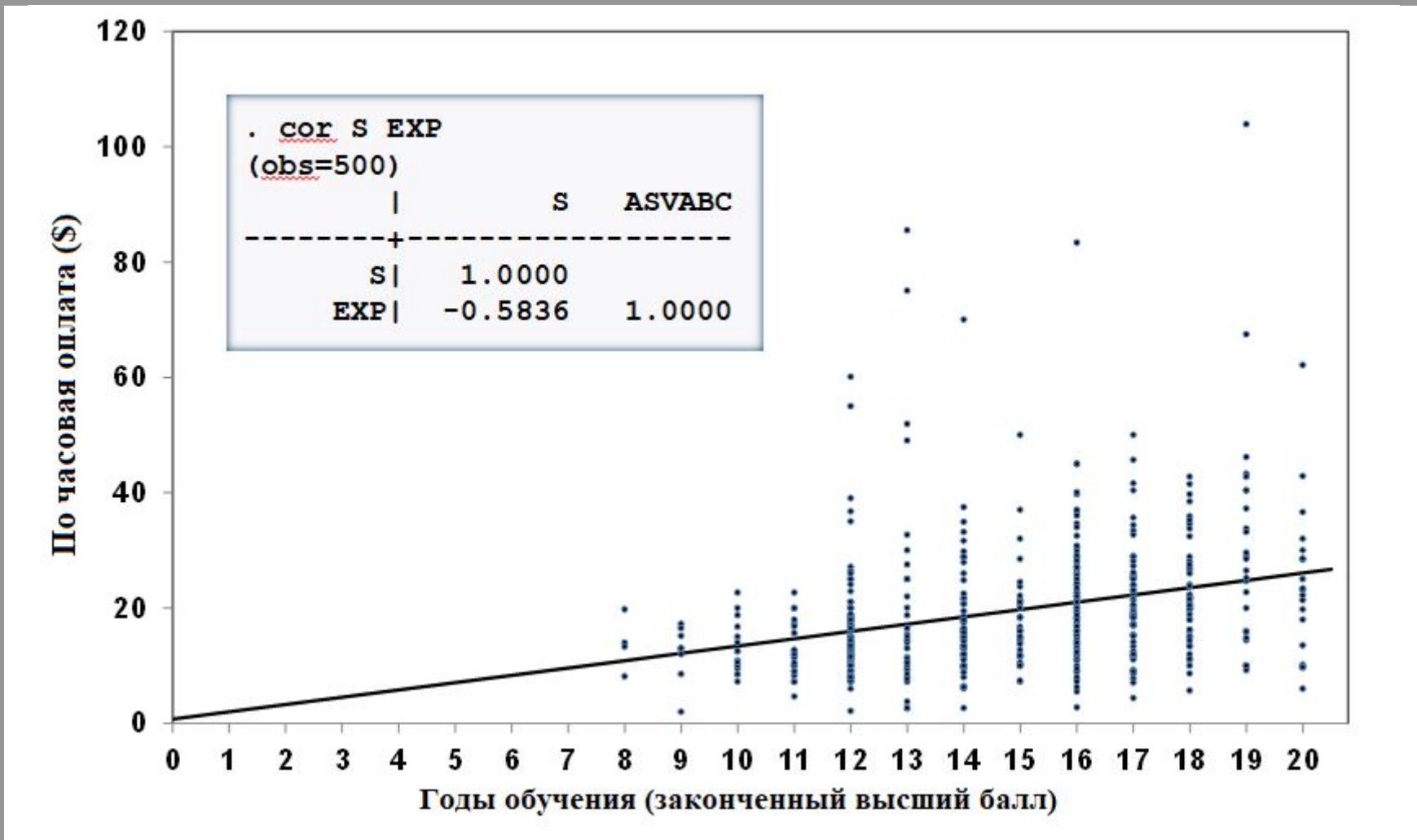
Простой график ввел бы в заблуждение

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА



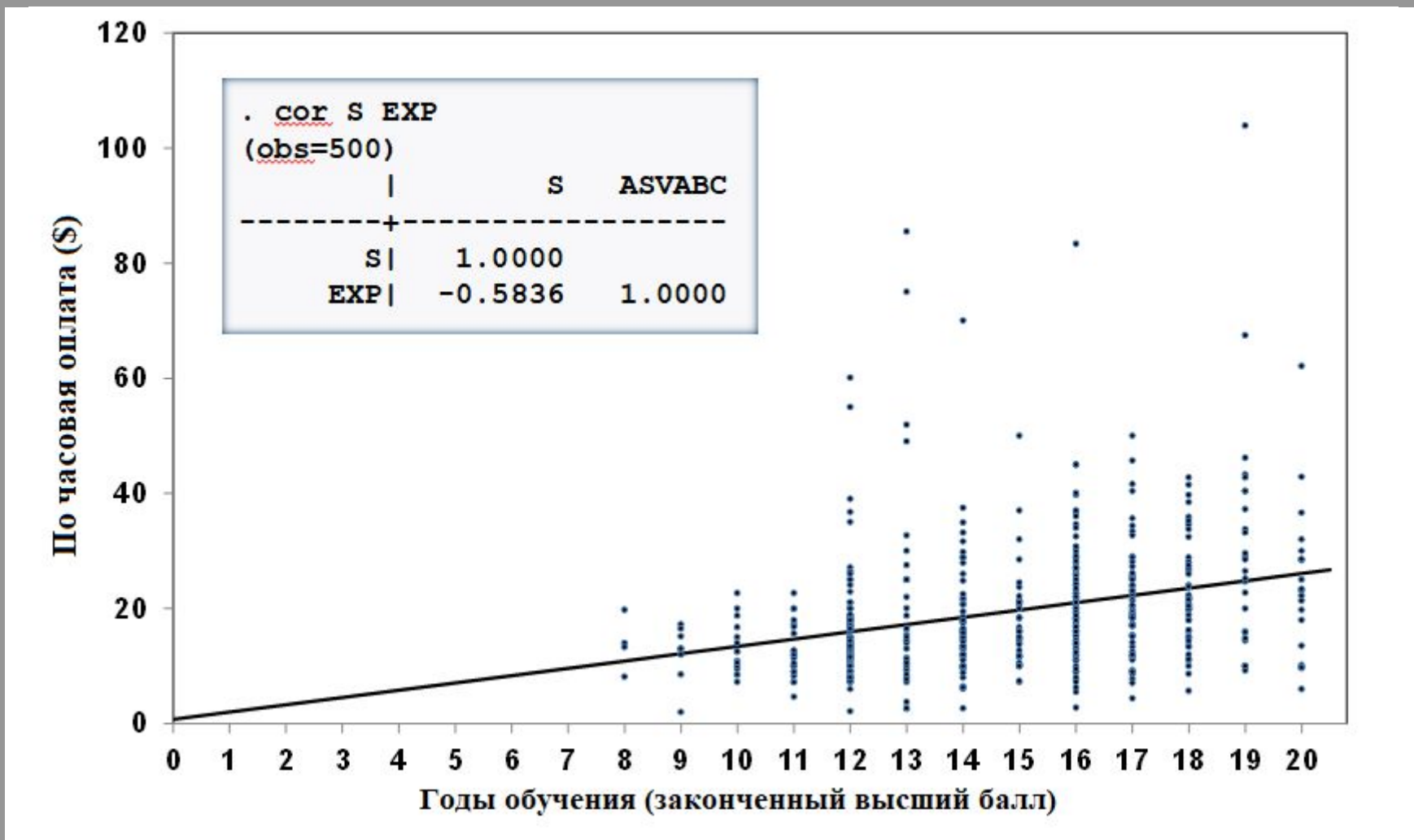
Обучение отрицательно коррелируется с опытом работы. График проваливается и нужно обратить на это внимание, и как следствие линия регресса недооценивает воздействие обучения на доход

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА



Мы исследуем искажение математически, когда мы дойдем до смещения из-за снятых переменных

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА



Чтобы устранить искажение, мы производим чистку обоих и ДОХОДОВ и S, их компонентов, связанных с EXP, и затем рисуем диаграмму разброса, используя очищенные переменные.

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА

```
. reg EARNINGS EXP
```

Source	SS	df	MS			
Model	8.36885807	1	8.36885807	Number of obs =	500	
Residual	70320.5974	498	141.206019	F(1, 498) =	0.06	
Total	70328.9662	499	140.939812	Prob > F =	0.8078	
				R-squared =	0.0001	
				Adj R-squared =	-0.0019	
				Root MSE =	11.883	

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EXP	-.0442828	.1818981	-0.24	0.808	-.4016651	.3130996
_cons	19.86614	1.287089	15.43	0.000	17.33735	22.39494

```
. predict EEARN, resid
```

Мы начинаем, возвращая ДОХОД на EXP, как показано выше. Остатки - часть ДОХОДА, которая не связана с ЭКСПОРТОМ, команда 'предсказать' - команда Stata для экономии остатков от нового регресса. Мы называем их EEARN.

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА

```
. reg S EXP
```

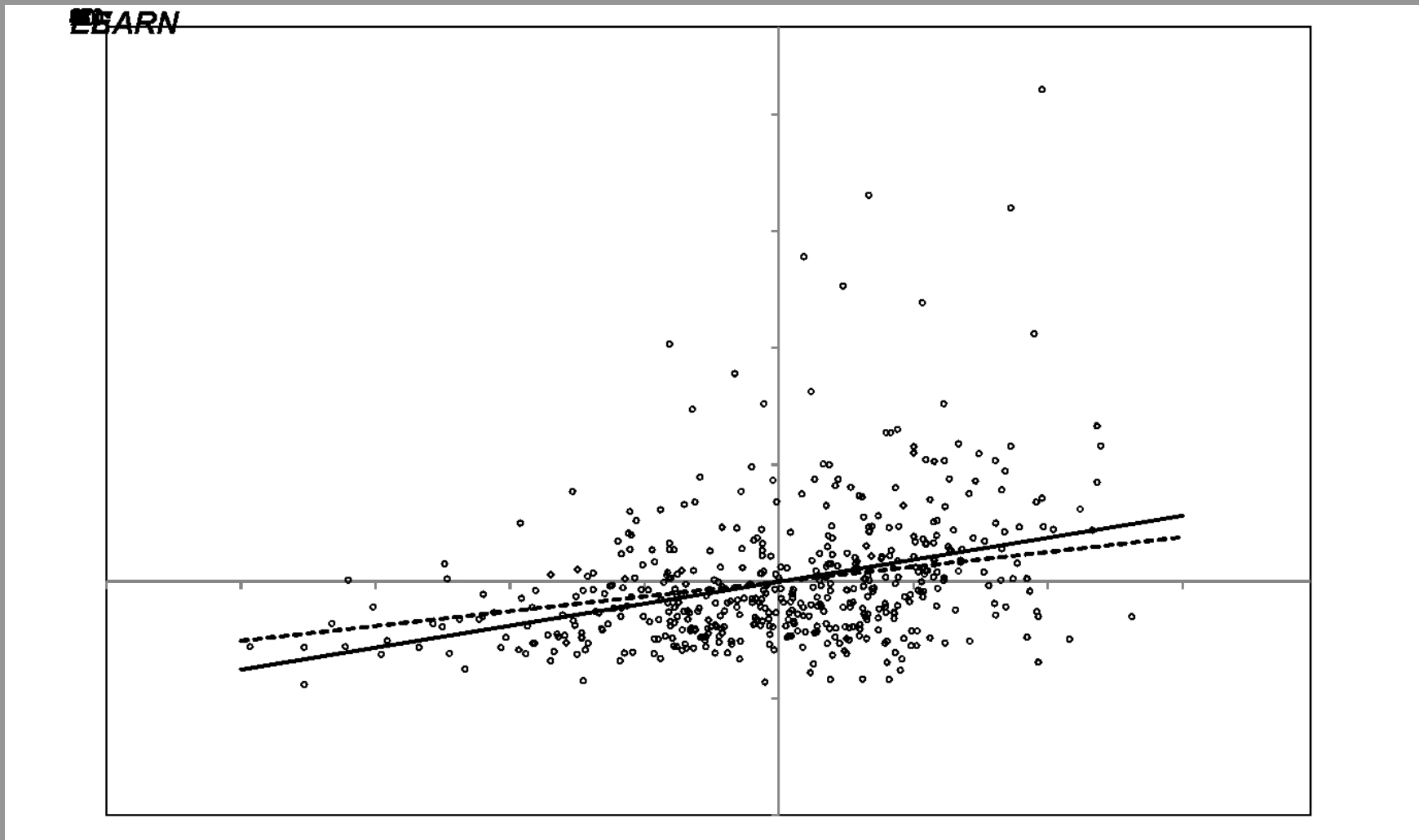
Source	SS	df	MS			
Model	1278.43322	1	1278.43322	Number of obs =	500	
Residual	2475.58878	498	4.9710618	F(1, 498) =	257.18	
				Prob > F =	0.0000	
				R-squared =	0.3406	
				Adj R-squared =	0.3392	
Total	3754.022	499	7.52309018	Root MSE =	2.2296	

S	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
EXP	-.5473191	.0341292	-16.04	0.000	-.6143741	-.4802641
_cons	18.39324	.241494	76.16	0.000	17.91877	18.86771

```
. predict ES, resid
```

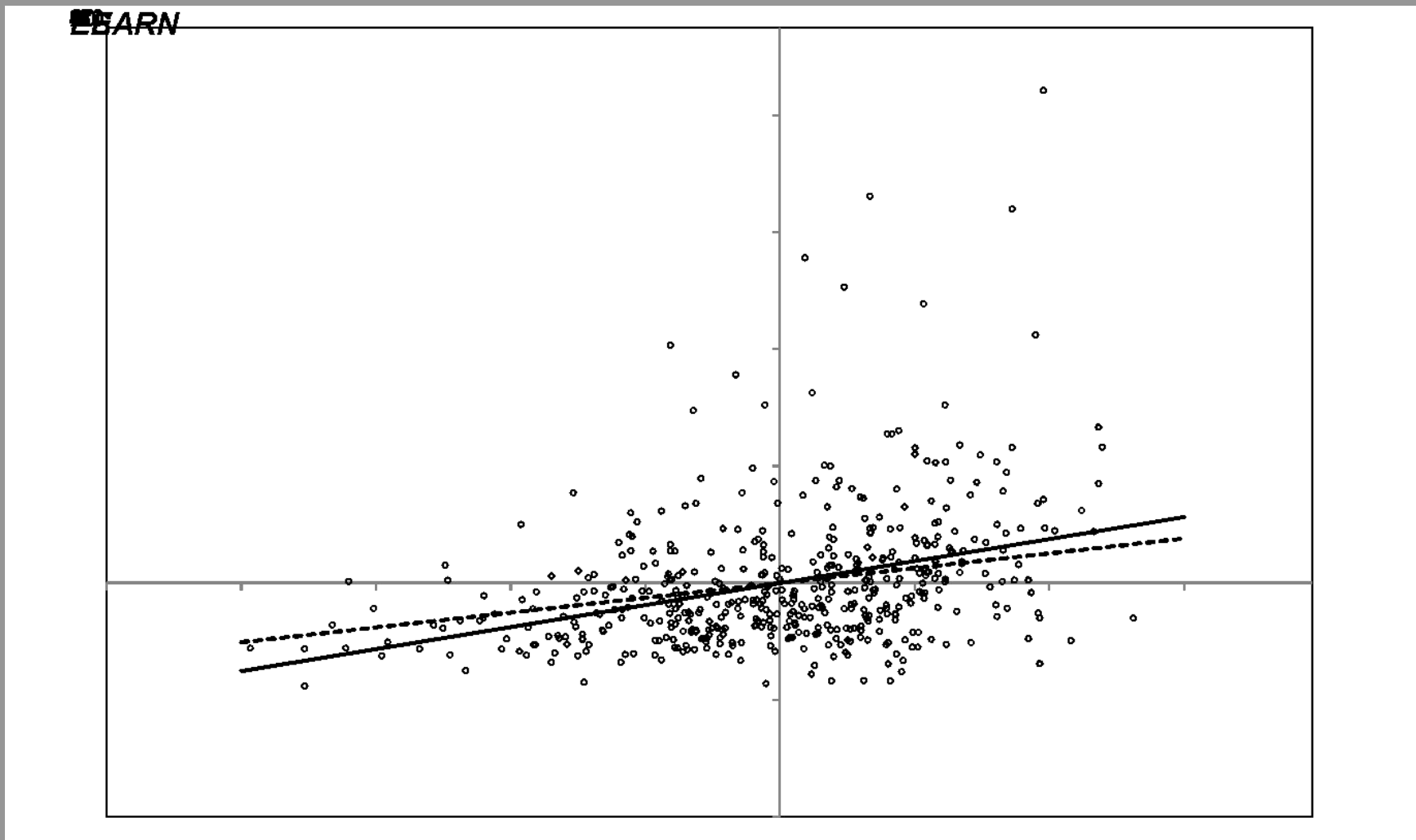
Мы делаем то же самое с S. Мы возвращаемся к EXP и сохраняем остатки как ES

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА



Теперь мы готовим EEARN на ES, и разброс - верное представление отношений, обоих с точки зрения наклона линии тенденции (твердая линия) и с точки зрения изменения о той линии.

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА



Как Вы ожидали бы, линия тенденции более крута, что в диаграмме разброса, которая не управляла для EXP (воспроизведенный здесь как пунктирная линия).

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА

```
. reg EEARN ES
```

Source	SS	df	MS				
Model	8727.05507	1	8727.05507	Number of obs = 500			
Residual	61593.5414	498	123.68181	F(1, 498) = 70.56			
				Prob > F = 0.0000			
				R-squared = 0.1241			
				Adj R-squared = 0.1223			
				Root MSE = 11.121			
EEARN	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]		
ES	1.877563	.2235186	8.40	0.000	1.438408	2.316719	
_cons	-1.32e-08	.4973566	-0.00	1.000	-.977176	.977176	

Вот регресс EEARN на ES.

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА

```
. reg EEARN ES
```

Source	SS	df	MS			
Model	8727.05507	1	8727.05507	Number of obs =	500	
Residual	61593.5414	498	123.68181	F(1, 498) =	70.56	
				Prob > F =	0.0000	
				R-squared =	0.1241	
				Adj R-squared =	0.1223	
Total	70320.5965	499	140.923039	Root MSE =	11.121	

EEARN	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ES	1.877563	.2235186	8.40	0.000	1.438408	2.316719
_cons	-1.32e-08	.4973566	-0.00	1.000	-.977176	.977176

From multiple regression:

```
. reg EARNINGS S EXP
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
S	1.877563	.2237434	8.39	0.000	1.437964	2.317163
EXP	.9833436	.2098457	4.69	0.000	.5710495	1.395638
_cons	-14.66833	4.288375	-3.42	0.001	-23.09391	-6.242752

Математическое доказательство, что работы техники требуют матричной алгебры. Мы будем довольствоваться, проверяя, что оценка наклонного коэффициента совпадает с в многократном регрессе.

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА

```
. reg EEARN ES
```

Source	SS	df	MS	Number of obs = 500		
Model	8727.05507	1	8727.05507	F(1, 498)	=	70.56
Residual	61593.5414	498	123.68181	Prob > F	=	0.0000
-----+-----				R-squared	=	0.1241
Total	70320.5965	499	140.923039	Adj R-squared	=	0.1223
-----+-----				Root MSE	=	11.121
EEARN	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ES	1.877563	.2235186	8.40	0.000	1.438408	2.316719
_cons	-1.32e-08	.4973566	-0.00	1.000	-.977176	.977176

From multiple regression:

```
. reg EARNINGS S EXP
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
S	1.877563	.2237434	8.39	0.000	1.437964	2.317163
EXP	.9833436	.2098457	4.69	0.000	.5710495	1.395638
_cons	-14.66833	4.288375	-3.42	0.001	-23.09391	-6.242752

Наконец, маленькое и не очень важный технический пункт. Вы, возможно, заметили, что стандартная ошибка и t статистическая величина действительно не совсем соответствуют. Причина этого состоит в том, что количество степеней свободы завышено 1 в регрессе остатков.

ГРАФИК ОТНОШЕНИЙ В МНОГОЧИСЛЕННОЙ МОДЕЛИ РЕГРЕССА

```
. reg EEARN ES
```

Source	SS	df	MS	Number of obs = 500		
Model	8727.05507	1	8727.05507	F(1, 498)	=	70.56
Residual	61593.5414	498	123.68181	Prob > F	=	0.0000
-----+-----				R-squared	=	0.1241
Total	70320.5965	499	140.923039	Adj R-squared	=	0.1223
-----+-----				Root MSE	=	11.121
EEARN	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ES	1.877563	.2235186	8.40	0.000	1.438408	2.316719
_cons	-1.32e-08	.4973566	-0.00	1.000	-.977176	.977176

From multiple regression:

```
. reg EARNINGS S EXP
```

EARNINGS	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
S	1.877563	.2237434	8.39	0.000	1.437964	2.317163
EXP	.9833436	.2098457	4.69	0.000	.5710495	1.395638
_cons	-14.66833	4.288375	-3.42	0.001	-23.09391	-6.242752

Тот регресс не делал поправку на то, что мы уже израсходовали 1 степень свободы в удалении EXP из модели.