

Penerapan Metode *Three-Stage Least Square* pada Persamaan Simultan Indeks Pembangunan Manusia dan Jumlah Penduduk Miskin

Fadila Nur Choirunnisa*, Nur Azizah Komara Rifai

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*fnurchoirunnisa@gmail.com, nur.azizah@unisba.ac.id

Abstract. Regression analysis is a method used to determine the influence between dependent variables and one or more independent variables. The OLS method is employed to estimate a single equation. In simultaneous equations, the OLS method cannot be utilized due to the correlation between independent variables and errors. Several OLS estimates are in simultaneous equations, including the Two-Stage Least Squares (2SLS) method and the Three-Stage Least Squares (3SLS) method. The 3SLS method combines the 2SLS method with the Seemingly Unrelated Regression (SUR) method. The first stage involves estimating the reduced-form equation using the OLS method. In the second stage, the results of this estimation are used to estimate the structural equation of the model. Finally, in the third stage, the SUR method is applied with Generalized Least Squares (GLS) estimation. This study uses the 3SLS method to simultaneous equations of the Human Development Index (HDI) and the number of poor populations. The estimation results for the HDI concurrent equation model indicate that population density has a positive and significant impact on HDI, and the poverty line has a positive and significant effect on HDI. In contrast, the simultaneous equation for the number of poor populations shows that the open unemployment rate (OUR) has a positive and significant effect on the number of poor populations. Therefore, there is no reciprocal relationship or simultaneous relationship between the HDI variable and the variable of the number of poor populations.

Keywords: *Human Development Index, Number of Poor Populations, Simultaneous Equations, Three-Stage Least Squares.*

Abstrak. Analisis regresi merupakan analisis yang digunakan untuk mengetahui pengaruh antara variabel tak bebas dengan satu atau lebih variabel bebas. Metode OLS digunakan untuk mengestimasi persamaan tunggal. Pada persamaan simultan metode OLS tidak dapat digunakan karena adanya korelasi antara variabel bebas dengan galat. Terdapat beberapa estimasi OLS dalam persamaan simultan yang dapat digunakan diantaranya metode *Two-Stage Least Square* (2SLS) dan metode *Three-Stage Least Squares* (3SLS). Metode 3SLS merupakan metode gabungan 2SLS dengan metode *Seemingly Unrelated Regression* (SUR). Tahap pertama estimasi persamaan reduksi dengan metode OLS, tahap kedua hasil estimasi tersebut digunakan untuk mengestimasi persamaan struktural dari model, kemudian tahap ketiga menggunakan metode SUR dengan estimasi *Generalized Least Square* (GLS). Penelitian ini menerapkan metode 3SLS pada persamaan simultan indeks pembangunan manusia (IPM) dan jumlah penduduk miskin. Hasil estimasi yang diperoleh pada model persamaan simultan IPM menunjukkan bahwa kepadatan penduduk berpengaruh positif dan signifikan terhadap IPM dan garis kemiskinan berpengaruh positif dan signifikan terhadap IPM. Sedangkan, pada persamaan simultan jumlah penduduk miskin menunjukkan bahwa tingkat pengangguran terbuka (TPT) berpengaruh positif dan signifikan terhadap jumlah penduduk miskin. Oleh karena itu, variabel IPM dan variabel jumlah penduduk miskin tidak ada hubungan timbal balik atau hubungan simultan.

Kata Kunci: *Indeks Pembangunan Manusia, Jumlah Penduduk Miskin, Persamaan Simultan, Three-Stage Least Square.*

A. Pendahuluan

Analisis yang digunakan untuk mengetahui pengaruh antara variabel bebas dengan satu atau lebih variabel bebas disebut analisis regresi. Berdasarkan jumlah variabel bebas, terdapat dua jenis analisis regresi yaitu analisis regresi sederhana (terdiri dari satu variabel bebas) dan analisis regresi berganda (lebih dari satu variabel bebas). Persamaan regresi diperoleh dengan mengestimasi parameter persamaan menggunakan metode Ordinary Least Square (OLS). Apabila terdapat dua atau lebih variabel tak bebas (Y) dan terdapat saling ketergantungan antar keduanya maka akan terjadi hubungan dua arah (hubungan timbal balik). Terjadinya hubungan dua arah pada model disebut dengan model persamaan simultan atau sistem persamaan simultan (Sugiyono, 2010).

Persamaan simultan memiliki lebih dari satu persamaan. Variabel tak bebas dalam persamaan simultan memiliki peran juga sebagai variabel bebas pada persamaan lainnya. Dalam persamaan simultan digunakan variabel tak bebas (variabel endogen) dan variabel bebas (variabel predetermined). Terdapat beberapa estimasi OLS dalam persamaan simultan yang dapat digunakan diantaranya Two-Stage Least Square (2SLS) dan Three-Stage Least Square (3SLS). Metode 3SLS terdiri dari tiga tahap, untuk tahap pertama dan kedua menggunakan tahapan dari metode 2SLS dan langkah ke tiga adalah menggunakan metode Seemingly Unrelated Regression (SUR).

Proses Pembangunan manusia di Indonesia diukur melalui Indeks Pembangunan Manusia (IPM). Antara tahun 2017 hingga 2021 IPM Provinsi Jawa Barat di peringkat 10 tanpa peningkatan. Salah satu faktor utama rendahnya IPM adalah peningkatan jumlah penduduk miskin. Hal tersebut menjadi hambatan dalam proses pembangunan dan menekankan pentingnya upaya pemberdayaan ekonomi, pendidikan, dan faktor lainnya untuk meningkatkan kualitas kehidupan dan mendorong pembangunan manusia.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Bagaimana model persamaan simultan pada IPM dan jumlah penduduk miskin dan variabel apa saja yang dapat mempengaruhi IPM dan jumlah penduduk miskin?”. Selanjutnya, berdasarkan rumusan masalah yang telah diketahui maka tujuan yang ingin dicapai adalah sebagai berikut.

1. Memperoleh model persamaan simultan pada IPM dan jumlah penduduk miskin.
2. Menentukan variabel-variabel yang mempengaruhi IPM dan jumlah penduduk miskin.

B. Metodologi Penelitian

Penelitian yang dilakukan menggunakan data sekunder dengan memperoleh data dari Badan Pusat Statistik Provinsi Jawa Barat. Data yang mencakup 27 Kabupaten dan Kota pada tahun 2021. Variabel yang digunakan dalam penelitian ini meliputi:

1. Variabel endogen
 Y_1 = Indeks Pembangunan Manusia (IPM)
 Y_2 = Jumlah Penduduk Miskin
2. Variabel predetermined
 X_1 = Kepadatan Penduduk
 X_2 = Tingkat Partisipasi Angkatan Kerja (TPAK)
 X_3 = Garis Kemiskinan
 X_4 = Umur Harapan Hidup
 X_5 = Tingkat Pengangguran Terbuka (TPT)

Tahapan analisis yang digunakan pada penelitian ini adalah sebagai berikut.

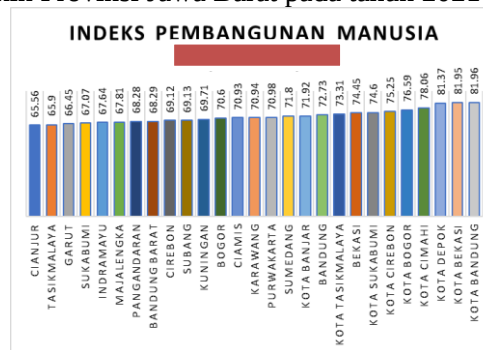
1. Menentukan model persamaan simultan IPM dan jumlah penduduk miskin.
2. Identifikasi model.
3. Melakukan *reduce form*.
4. Melakukan uji Hausman.
5. Mengestimasi persamaan simultan dengan metode 3SLS.

6. Melakukan uji simultan dan uji parsial.
7. Menentukan koefisien determinasi.
8. Melakukan pengujian asumsi klasik.

C. Hasil Penelitian dan Pembahasan

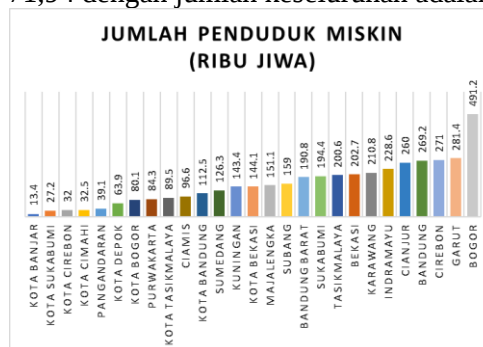
Analisis Deskriptif

Berikut adalah data yang digunakan dalam penelitian ini terdiri dari 27 Kabupaten/ Kota di Provinsi Jawa Barat pada tahun 2021. Berikut merupakan grafik yang disajikan mengenai IPM dan jumlah penduduk miskin Provinsi Jawa Barat pada tahun 2021 per Kabupaten/ Kota.



Gambar 1. IPM Provinsi Jawa Barat Tahun 2021

Berdasarkan Gambar 1. menunjukkan nilai IPM terendah adalah sebesar 65,56 di Kota Cianjur, sementara nilai tertinggi adalah sebesar 81,96 di Kota Bandung. Diperoleh rata-rata dari nilai IPM yaitu sebesar 71,94 dengan jumlah keseluruhan adalah sebesar 1942,40.



Gambar 2. Jumlah Penduduk Miskin Provinsi Jawa Barat Tahun 2021

Berdasarkan Gambar 3. menunjukkan nilai terendah dari jumlah penduduk miskin adalah sebesar 13,4 di Kota Banjar dan nilai tertinggi sebesar 491,2 di Kabupaten Bogor. Diperoleh rata-rata jumlah penduduk miskin sebesar 155,396 dengan jumlah keseluruhan adalah sebesar 4195,7.

Model Persamaan Simultan IPM Dan Jumlah Penduduk Miskin

Penyusunan bentuk persamaan IPM dan jumlah penduduk miskin didasarkan pada hasil penelitian sebelumnya yang telah dijelaskan pada latar belakang. Selain itu, penelitian yang dilakukan oleh Hutabarat *et al* (2021) menunjukkan bahwa umur harapan hidup berpengaruh negatif secara signifikan terhadap jumlah penduduk miskin di Sumatera Utara. Tujuan dari penyusunan persamaan ini adalah untuk menggambarkan hubungan antara IPM dan jumlah penduduk miskin serta memahami faktor-faktor yang berpengaruh.

Berikut merupakan spesifikasi model persamaan simultan yang digunakan adalah

sebagai berikut.

$$Y_1 = \beta_{10} + \beta_{12}Y_2 + \gamma_{11}X_1 + \gamma_{12}X_2 + \gamma_{13}X_3 + u_1 \quad \dots(3.1)$$

$$Y_2 = \beta_{20} + \beta_{21}Y_1 + \gamma_{22}X_2 + \gamma_{24}X_4 + \gamma_{25}X_5 + u_2 \quad \dots(3.2)$$

Identifikasi Model

Identifikasi model dalam persamaan simultan penting untuk memperoleh estimasi parameter. Model persamaannya telah diperoleh terdiri dari persamaan IPM dan persamaan jumlah penduduk miskin. Setelah itu adalah melakukan identifikasi model untuk mengetahui metode 3SLS dapat digunakan.

Tabel 1. Hasil Identifikasi Model

Persamaan	k	k^*	n	$k - k^* > n - 1$	Teridentifikasi
Y_1	5	3	2	Ya	<i>overidentified</i>
Y_2	5	3	2	Ya	<i>overidentified</i>

Berdasarkan Tabel 1. diatas dapat dilihat bahwa identifikasi persamaan dengan menggunakan *orde condition* dari kedua persamaan teridentifikasi *overidentified*. Sehingga metode estimasi dengan menggunakan metode 3SLS dapat digunakan.

Reduce Form

Reduce form dilakukan untuk menyajikan variabel-variabel endogen sebagai fungsi dari variabel-variabel *predetermined* yang akan dijadikan sebagai variabel instrumental (Wulandari, 2010). Persamaan *reduce form* diperoleh dari mengubah persamaan model persamaan simultan menjadi persamaan *reduce form*, maka diperoleh hasilnya sebagai berikut.

$$Y_1 = \pi_0 + \pi_1X_1 + \pi_2X_2 + \pi_3X_3 + \pi_4X_4 + \pi_5X_5 + u_1 \quad \dots(3.3)$$

$$Y_2 = \pi_6 + \pi_7X_1 + \pi_8X_2 + \pi_9X_3 + \pi_{10}X_4 + \pi_{11}X_5 + u_2 \quad \dots(3.4)$$

Uji Hausman

Uji Hausman dilakukan untuk mengetahui variabel *predetermined* bersifat endogen. Metode 3SLS dapat digunakan jika terjadi endogenitas. Endogenitas terjadi ketika variabel bebas saling terkait dengan variabel endogen dalam model regresi, sehingga menyebabkan bias dan ketidakseimbangan estimasi model. Maka hasil uji Hausman adalah sebagai berikut.

Tabel 2. Hasil Uji Hausman

<i>Hausman Test</i>	t_{hitung}	$p-value$
Y_1	5,10495	5×10^{-5}
Y_2	11,82941	1×10^{-10}

Berdasarkan **Tabel 2.** hasil uji Hausman yang diperoleh dari persamaan Y_1 adalah nilai $p-value = 5 \times 10^{-5} < 0,05$ maka H_0 ditolak dan persamaan Y_2 diperoleh hasil uji Hausman dengan nilai $p-value = 1 \times 10^{-10} < 0,05$ maka H_0 ditolak. Artinya, pada kedua persamaan terdapat simultanitas. Sehingga penggunaan metode 3SLS tersebut tepat karena kedua persamaan tersebut endogen.

Metode Three-Stage Least Square

Hasil estimasi persamaan simultan dengan metode 3SLS yang menggunakan *software* R-Studio pada persamaan IPM dan jumlah penduduk miskin, sebagai berikut.

Tabel 3. Hasil estimasi metode 3SLS persamaan IPM

Persamaan Y_1 (IPM)					
Variabel	Koefisien	Standar Error	t_{hitung}	$p-value$	F_{hitung}
Intercept	65,83403	8,62864	7,6297	$1,3 \times 10^{-7}$	32,19705 $p-value = 6 \times 10^{-9}$
Y_2^*	-0,13709	0,17048	-0,8041	0,4299	
X_1	0,00068	$1,5 \times 10^{-4}$	4,3722	$2,4 \times 10^{-4}$	
X_2	-0,00720	0,11908	-0,0605	0,9523	
X_3	0,00001	5×10^{-5}	2,3026	0,0311	

Tabel 4. Hasil estimasi 3SLS persamaan jumlah penduduk miskin

Persamaan Y_2 (Jumlah Penduduk Miskin)					
Variabel	Koefisien	Standar Error	t_{hitung}	$p-value$	F_{hitung}
Intercept	32,16347	56,64569	0,5678	0,5759	3,89689 $p-value = 0,02005$
Y_1	-0,49207	0,33260	-1,4795	0,1532	
X_2	0,50109	0,30694	1,6325	0,1168	
X_4	-0,41497	1,12442	-0,3691	0,7156	
X_5	1,32130	0,48561	2,7209	0,0125	

Berdasarkan secara keseluruhan maka model yang diperoleh dari Tabel 3 dan Tabel 4 sebagai berikut.

$$\hat{Y}_1 = 65,83403 - 0,13709Y_2^* + 0,00068X_1 - 0,00720X_2 + 0,00001X_3 \quad \dots(3.5)$$

$$R^2 = 0,8541$$

$$\hat{Y}_2^* = 32,16347 - 0,49207Y_1 + 0,50109X_2 - 0,41497X_4 + 1,32130X_5 \quad \dots(3.6)$$

$$R^2 = 0,4147$$

Berdasarkan persamaan (3.5) diperoleh bahwa nilai koefisien determinasi persamaan Y_1 (IPM) dengan menggunakan 3SLS adalah sebesar 0,8541 atau 85,41%, sisanya dipengaruhi oleh variabel yang lain. Sedangkan, persamaan (3.6) diperoleh bahwa nilai koefisien determinasi persamaan Y_2 (jumlah penduduk miskin) adalah sebesar 0,4147 atau 41,47%, sisanya dipengaruhi oleh variabel lain.

Berdasarkan model yang diperoleh dari persamaan (3.5) variabel yang signifikan berpengaruh terhadap IPM dengan nilai $\alpha = 0,05$ dimana $p-value < \alpha$, sehingga kepadatan penduduk dan garis kemiskinan berpengaruh positif terhadap IPM. Sedangkan model yang diperoleh dari persamaan (3.6) variabel yang signifikan berpengaruh terhadap jumlah penduduk miskin dengan nilai $\alpha = 0,05$ dimana $p-value < \alpha$, sehingga TPT berpengaruh positif terhadap jumlah penduduk miskin.

Uji Asumsi Klasik

Uji normalitas

Uji normalitas digunakan untuk mengetahui galat berdistribusi normal atau tidak. Pada uji normalitas ini menggunakan uji *Shapiro-Wilk*, karena data yang digunakan sebanyak 27. Maka diperoleh hasil uji normalitas sebagai berikut:

Tabel 5. Hasil Uji Normalitas

Persamaan	<i>Shapiro-Wilk</i>	$p-value$
Y_1	0,96076	0,3845
Y_2	0,97907	0,8407

Berdasarkan Tabel 5. diperoleh dari persamaan Y_1 (IPM) dengan nilai $p\text{-value} = 0,3845$ dimana nilai $p\text{-value} > \alpha$ maka H_0 diterima sehingga persamaan Y_1 (IPM) galat berdistribusi normal. Sedangkan dari persamaan Y_2 (jumlah penduduk miskin) diperoleh bahwa nilai $p\text{-value} = 0,8407$ dimana nilai $p\text{-value} > \alpha$ maka H_0 diterima sehingga persamaan Y_2 (jumlah penduduk miskin) galatnya berdistribusi normal.

Uji Heteroskedastisitas

Uji heteroskedastisitas untuk mengetahui terdapat ketidaksamaan varian atau tidak. Pengujian heteroskedastisitas dilakukan menggunakan uji *Breusch-Pagan*.

Tabel 6. Hasil Uji Heteroskedastisitas

Persamaan	<i>Breusch-Pagan test</i>	<i>p-value</i>
Y_1	2,3031	0,6802
Y_2	0,98985	0,9113

Berdasarkan Tabel 6. diperoleh Persamaan Y_1 diperoleh bahwa nilai $p\text{-value} = 0,6802$ dengan $p\text{-value} > \alpha$ maka H_0 diterima sehingga persamaan Y_1 (IPM) tidak terjadi heteroskedastisitas. Sedangkan, persamaan Y_2 (jumlah penduduk miskin) diperoleh bahwa nilai $p\text{-value} = 0,9113$ dengan $p\text{-value} > \alpha$ maka H_0 diterima sehingga persamaan Y_2 (jumlah penduduk miskin) tidak terjadi heteroskedastisitas.

Multikolinearitas

Pengujian ini dilakukan untuk melihat korelasi antar variabel *predetermined* atau variabel endogen. Pengujian multikolinearitas dilakukan dengan menggunakan nilai *Variance Inflation Factor* (VIF).

Tabel 7. Hasil Multikolinearitas IPM

Variabel	VIF	Ketentuan
Jumlah Penduduk Miskin	1,2483	Nilai VIF < 10
Kepadatan Penduduk	2,8328	Nilai VIF < 10
TPAK	1,2032	Nilai VIF < 10
Garis Kemiskinan	2,2735	Nilai VIF < 10

Tabel 8. Hasil Multikolinearitas Jumlah Penduduk Miskin

Variabel	VIF	Ketentuan
IPM	3,7493	Nilai VIF < 10
TPAK	2,3591	Nilai VIF < 10
Umur harapan hidup	4,2552	Nilai VIF < 10
TPT	2,7955	Nilai VIF < 10

Berdasarkan Tabel 7. dan Tabel 8. bahwa nilai VIF < 10 pada semua variabel sehingga dapat disimpulkan tidak terjadi multikolinearitas pada persamaan IPM dan jumlah penduduk miskin.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

- Hasil estimasi menggunakan metode *Three-Stage Least Square* (3SLS) menghasilkan model persamaan simultan sebagai berikut.

$$\hat{Y}_1 = 65,83403 - 0,13709 \text{ Jumlah Penduduk Miskin} + 0,00068 \text{ Kepadatan Penduduk} - 0,00720 \text{ Tingkat Partisipasi Angkatan Kerja} + 0,00001 \text{ Garis}$$

Kemiskinan

$$\hat{Y}_2^* = 32,16347 - 0,49207 \text{ IPM} + 0,50109 \text{ Tingkat Partisipasi Angkatan Kerja} - 0,41497 \text{ Umur Harapan Hidup} + 1,32130 \text{ Tingkat Pengangguran Terbuka}$$

2. Hasil pengujian diperoleh bahwa IPM dipengaruhi kepadatan penduduk dengan nilai *p-value* sebesar 0,00024 dan nilai koefisien sebesar 0,00068 yang menunjukkan bahwa kepadatan penduduk berpengaruh positif dan signifikan terhadap IPM, garis kemiskinan dengan nilai *p-value* 0,0311 dan nilai koefisien 0,00001 yang menunjukkan bahwa garis kemiskinan berpengaruh positif dan signifikan terhadap IPM. Sedangkan, jumlah penduduk miskin dipengaruhi oleh tingkat pengangguran terbuka (TPT) dengan nilai *p-value* sebesar 0,0125 dan nilai koefisien 1,32130 yang menunjukkan bahwa tingkat pengangguran terbuka (TPT) berpengaruh positif terhadap jumlah penduduk miskin.

Acknowledge

Penulis mengucapkan terima kasih kepada semua pihak yang telah membantu kelancaran dalam melaksanakan penelitian ini.

Daftar Pustaka

- [1] Badan Pusat Statistik Provinsi Jawa Barat (<https://jabar.bps.go.id/>, diakses Maret 2023).
- [2] Drapper, N. dan H. Smith.(1992). Analisis Regresi Terapan (Edisi Kedua). Jakarta: Alih Bahasa oleh penerbit PT. Gramedia Pustaka Utama.
- [3] Greene, W. H. (1993). *Econometric Analysis Fifth Edition*, New Jersey: Prentice Hall.
- [4] Gujarati D. *Ekonometria Dasar*, Terjemahan: Drs. Ak. Sumarno Zain, MBA. Jakarta: Erlangga; 2003.
- [5] Hausman, J. A. (1978) Specification Tests in Econometrics. *Econometrica*. 46:1251-1271.
- [6] Hutabarat, Desty. (2018). Pengaruh Angka Harapan Hidup, Rata-Rata Lama Sekolah, Pengeluaran Rill Perkapita, Pertumbuhan Ekonomi dan Pengangguran terhadap Jumlah Penduduk Miskin di Provinsi Sumatera Utara. Skripsi. Sumatera Utara: Jurusan Ekonomi Pembangunan, Fakultas Ekonomi Bisnis, Universitas Sumatera Utara.
- [7] Misno dan Sulistianingsih, E. (2019). Estimasi Model Persamaan Simultan dengan Metode Two Stage Least Square (2SLS). *Buletin Ilmiah Math, Stat, dan Terapan (Bimaster)*, 653-658.
- [8] Wulandari, Septi. 2010. Penerapan Metode 2SLS (Two Stage Least Square) pada Model Persamaan Simultan Untuk Persamaan Pendapatan Nasional dan Peredaran Uang. Skripsi. Tidak Diterbitkan. Program Pasca Sarjana Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Sebelas Maret: Surakarta.
- [9] Devila Mustika Prancisca, & Darwis, S. (2021). Prediksi Sisa Umur Bearing Menggunakan Regresi Eksponensial. *Jurnal Riset Statistika*, 1(2), 107–116. <https://doi.org/10.29313/jrs.v1i2.434>
- [10] Ratih Nurfitri, & Teti Sofia Yanti. (2023). Pemodelan Umur Harapan Hidup di Jabar Tahun 2021 Menggunakan Spatial Durbin Model. *Jurnal Riset Statistika*, 137–146. <https://doi.org/10.29313/jrs.v3i2.3023>

- [11] Sahwa Chanigia Viqri, Z., & Kurniati, E. (2023). *Perbandingan Penerapan Metode Fuzzy Time Series Model Chen-Hsu dan Model Lee dalam Memprediksi Kurs Rupiah terhadap Dolar Amerika*. 1(1), 19–26. <https://doi.org/10.29313/datamath.v1i1.12>